Environmental Bisimulations for Delimited-Control Operators

Dariusz Biernacki, Sergueï Lenglet

Environmental Bisimulations for Delimited-Control Operators

Dariusz Biernacki[1] and Sergueï Lenglet[2]

[1] Institute of Computer Science, University of Wrocław
[2] LORIA, Université de Lorraine

Abstract. We present a theory of environmental bisimilarity for the delimited-control operators *shift* and *reset*. We consider two different notions of contextual equivalence: one that does not require the presence of a top-level control delimiter when executing tested terms, and another one, fully compatible with the original CPS semantics of shift and reset, that does. For each of them, we develop sound and complete environmental bisimilarities, and we discuss up-to techniques.

1 Introduction

Control operators for delimited continuations [7, 9] provide elegant means for expressing advanced control mechanisms [7, 11]. Moreover, they play a fundamental role in the semantics of computational effects [10], normalization by evaluation [2] and as a crucial refinement of abortive control operators such as *callcc* [9, 20]. Of special interest are the control operators *shift* and *reset* [7] due to their origins in continuation-passing style (CPS) and their connection with computational monads – as demonstrated by Filinski [10], shift and reset can express in direct style arbitrary computational effects, such as mutable state, exceptions, etc. Operationally, the control delimiter reset delimits the current continuation and the control operator shift abstracts the current delimited continuation as a first class value that when resumed is composed with the then-current continuation.

Because of the complex nature of control effects, it can be difficult to determine if two programs that use shift and reset are equivalent (i.e., behave in the same way) or not. *Contextual equivalence* [16] is widely considered as the most natural equivalence on terms in languages similar to the λ-calculus. Roughly, two terms are contextually equivalent if we cannot tell them apart when they are executed within any context. The latter quantification over contexts makes this relation hard to use in practice, so we usually look for simpler characterizations of contextual equivalence, such as coinductively defined *bisimilarities*.

In our previous work, we defined *applicative* [4] and *normal form* [5] bisimilarities for shift and reset. Applicative bisimilarity characterizes contextual equivalence, but still quantifies over some contexts to relate terms (e.g., λ-abstractions are applied to the same arbitrary argument). As a result, some equivalences remain quite difficult to prove. In contrast, normal form bisimilarity does not contain any quantification over contexts or arguments in its definition: the tested terms are reduced to normal forms, which are then decomposed in bisimilar subterms. Consequently, proofs of equivalence are usually simpler than with applicative bisimilarity, and they can be simplified even further with *up-to techniques*. However, normal form bisimilarity is not *complete*, i.e., there exists contextually equivalent terms which are not normal form bisimilar.

Environmental bisimilarity [18] is a different kind of behavioral equivalence which in terms of strength and practicality can be situated in between applicative and normal form bisimilarities. It has originally been proposed in [22] and has been since defined in various higher-order languages (see, e.g., [19, 21, 17]). Like applicative bisimilarity, it uses some particular contexts to test terms, except that the testing contexts are built from an environment, which represents the knowledge built so far by an outside observer. Environmental bisimilarity usually characterizes contextual equivalence, but is harder to establish than applicative bisimilarity. Nonetheless, like with normal form bisimilarity, one can define powerful up-to techniques [18] to simplify the equivalence proofs. Besides, the authors of [14] argue that the additional complexity of environmental bisimilarity is necessary to handle more realistic features, like local state or exceptions.

In the quest for a powerful enough (i.e., as discriminative as contextual equivalence) yet easy-to-use equivalence for delimited control, we study in this paper the environmental theory of a calculus with shift and reset. More precisely, we consider two semantics for shift and reset: the original one [3], where terms are executed within a top-level reset, and a more relaxed semantics where this requirement is lifted. The latter is commonly used in implementations of shift and reset [8, 10] as well as in some studies of these operators [1, 12], including our previous work [4, 5]. So far, the behavioral theory of shift and reset with the original semantics has not been studied. Firstly, we define environmental bisimilarity for the relaxed semantics and study its properties; especially we discuss the problems raised by delimited control for the definition of bisimulation up to context, one of the most powerful up-to techniques. Secondly, we propose the first behavioral theory for the original semantics, and we pinpoint the differences between the equivalences of the two semantics. In particular, we show that the environmental bisimilarity for the original semantics is complete w.r.t. the axiomatization of shift and reset of [13], which is not the case for the relaxed semantics, as already proved in [4] for applicative bisimilarity.

In summary, we make the following contributions in this paper.

- We show that environmental bisimilarity can be defined for a calculus with delimited control, for which we consider two different semantics. In each case, the defined bisimilarity equals contextual equivalence.
- For the relaxed semantics, we explain how to handle *stuck terms*, i.e., terms where a capture cannot go through because of the lack of an outermost reset.
- We discuss the limits of the usual up-to techniques in the case of delimited control.
- For the original semantics, we define a contextual equivalence, and a corresponding environmental bisimilarity. Proving soundness of the bisimilarity w.r.t. contextual equivalence requires significant changes from the usual

soundness proof scheme. We discuss how environmental bisimilarity is easier to adapt than applicative bisimilarity.
– We give examples illustrating the differences between the two semantics.

The rest of the paper is organized as follows: in Section 2, we present the calculus $\lambda_{\mathcal{S}}$ used in this paper, and recall some results, including the axiomatization of [13]. We develop an environmental theory for the relaxed semantics in Section 3, and for the original semantics in Section 4. We conclude in Section 5, and the appendices contain the characterization proofs omitted from the main text.

2 The Calculus $\lambda_{\mathcal{S}}$

2.1 Syntax

The language $\lambda_{\mathcal{S}}$ extends the call-by-value λ-calculus with the delimited-control operators *shift* and *reset* [7]. We assume we have a set of term variables, ranged over by x, y, z, and k. We use k for term variables representing a continuation (e.g., when bound with a shift), while x, y, and z stand for any values; we believe such distinction helps to understand examples and reduction rules. The syntax of terms is given by the following grammar:

$$\text{Terms: } t ::= x \mid \lambda x.t \mid t\,t \mid \mathcal{S}k.t \mid \langle t \rangle$$

Values, ranged over by v, are terms of the form $\lambda x.t$. The operator shift ($\mathcal{S}k.t$) is a capture operator, the extent of which is determined by the delimiter reset ($\langle\cdot\rangle$). A λ-abstraction $\lambda x.t$ binds x in t and a shift construct $\mathcal{S}k.t$ binds k in t; terms are equated up to α-conversion of their bound variables. The set of free variables of t is written $\mathsf{fv}(t)$; a term t is *closed* if $\mathsf{fv}(t) = \emptyset$.

We distinguish several kinds of contexts, represented outside-in, as follows:

$$\begin{array}{ll} \text{Pure contexts:} & E ::= \square \mid v\,E \mid E\,t \\ \text{Evaluation contexts:} & F ::= \square \mid v\,F \mid F\,t \mid \langle F \rangle \\ \text{Contexts:} & C ::= \square \mid \lambda x.C \mid t\,C \mid C\,t \mid \mathcal{S}k.C \mid \langle C \rangle \end{array}$$

Regular contexts are ranged over by C. The pure evaluation contexts[3] (abbreviated as pure contexts), ranged over by E, represent delimited continuations and can be captured by shift. The call-by-value evaluation contexts, ranged over by F, represent arbitrary continuations and encode the chosen reduction strategy. Filling a context C (respectively E, F) with a term t produces a term, written $C[t]$ (respectively $E[t]$, $F[t]$); the free variables of t may be captured in the process. We extend the notion of free variables to contexts (with $\mathsf{fv}(\square) = \emptyset$), and we say a context C (respectively E, F) is *closed* if $\mathsf{fv}(C) = \emptyset$ (respectively $\mathsf{fv}(E) = \emptyset$, $\mathsf{fv}(F) = \emptyset$).

[3] This terminology comes from Kameyama (e.g., in [13]).

2.2 Reduction Semantics

The call-by-value reduction semantics of $\lambda_{\mathcal{S}}$ is defined as follows, where $t\{v/x\}$ is the usual capture-avoiding substitution of v for x in t:

$$
\begin{array}{ll}
(\beta_v) & F[(\lambda x.t)\ v] \rightarrow_{\mathrm{v}} F[t\{v/x\}] \\
(\mathit{shift}) & F[\langle E[\mathcal{S}k.t]\rangle] \rightarrow_{\mathrm{v}} F[\langle t\{\lambda x.\langle E[x]\rangle/k\}\rangle] \text{ with } x \notin \mathsf{fv}(E) \\
(\mathit{reset}) & F[\langle v\rangle] \rightarrow_{\mathrm{v}} F[v]
\end{array}
$$

The term $(\lambda x.t)\ v$ is the usual call-by-value redex for β-reduction (rule (β_v)). The operator $\mathcal{S}k.t$ captures its surrounding context E up to the dynamically nearest enclosing reset, and substitutes $\lambda x.\langle E[x]\rangle$ for k in t (rule (*shift*)). If a reset is enclosing a value, then it has no purpose as a delimiter for a potential capture, and it can be safely removed (rule (*reset*)). All these reductions may occur within a metalevel context F, so the reduction rules specify both the notion of reduction and the chosen call-by-value evaluation strategy that is encoded in the grammar of the evaluation contexts. Furthermore, the reduction relation $\rightarrow_{\mathrm{v}}$ is compatible with evaluation contexts F, i.e., $F[t] \rightarrow_{\mathrm{v}} F[t']$ whenever $t \rightarrow_{\mathrm{v}} t'$.

There exist terms which are not values and which cannot be reduced any further; these are called *stuck terms*.

Definition 1. *A term t is stuck if t is not a value and $t \not\rightarrow_{\mathrm{v}}$.*

For example, the term $E[\mathcal{S}k.t]$ is stuck because there is no enclosing reset; the capture of E by the shift operator cannot be triggered.

Lemma 1. *A closed term t is stuck iff $t = E[\mathcal{S}k.t']$ for some E, k, and t'.*

Definition 2. *A term t is a normal form if t is a value or a stuck term.*

We call *redexes* (ranged over by r) terms of the form $(\lambda x.t)\ v$, $\langle E[\mathcal{S}k.t]\rangle$, and $\langle v\rangle$. Thanks to the following unique-decomposition property, the reduction relation $\rightarrow_{\mathrm{v}}$ is deterministic.

Lemma 2. *For all closed terms t, either t is a normal form, or there exist a unique redex r and a unique context F such that $t = F[r]$.*

Finally, we write $\rightarrow_{\mathrm{v}}^{*}$ for the transitive and reflexive closure of $\rightarrow_{\mathrm{v}}$, and we define the evaluation relation of $\lambda_{\mathcal{S}}$ as follows.

Definition 3. *We write $t \Downarrow_{\mathrm{v}} t'$ if $t \rightarrow_{\mathrm{v}}^{*} t'$ and $t' \not\rightarrow_{\mathrm{v}}$.*

The result of the evaluation of a closed term, if it exists, is a normal form. If a term t admits an infinite reduction sequence, we say it *diverges*, written $t \Uparrow_{\mathrm{v}}$. Henceforth, we use $\Omega = (\lambda x.x\ x)\ (\lambda x.x\ x)$ as an example of such a term.

2.3 CPS Equivalence

In [13], the authors propose an equational theory of shift and reset based on CPS [7]. The idea is to relate terms that have $\beta\eta$-convertible CPS translations.

Definition 4. *Terms t_0 and t_1 are CPS equivalent, written $t_0 \equiv t_1$, if their CPS translations are $\beta\eta$-convertible.*

Kameyama and Hasegawa propose eight axioms in [13] to characterize CPS equivalence: two terms are CPS equivalent iff one can derive their equality using the equations below. Note that the axioms are defined on open terms, and suppose variables as values.

$$
\begin{array}{rlrl}
(\lambda x.t)\, v & =_{\text{KH}} t\{v/x\} & (\lambda x.E[x])\, t & =_{\text{KH}} E[t] \text{ if } x \notin \mathsf{fv}(E) \\
\langle E[\mathcal{S}k.t] \rangle & =_{\text{KH}} \langle t\{\lambda x.\langle E[x]\rangle/k\}\rangle & \langle (\lambda x.t_0)\, \langle t_1 \rangle \rangle & =_{\text{KH}} (\lambda x.\langle t_0 \rangle)\, \langle t_1 \rangle \\
\langle v \rangle & =_{\text{KH}} v & \mathcal{S}k.\langle t \rangle & =_{\text{KH}} \mathcal{S}k.t \\
\lambda x.v\, x & =_{\text{KH}} v \text{ if } x \notin \mathsf{fv}(v) & \mathcal{S}k.k\, t & =_{\text{KH}} t \text{ if } k \notin \mathsf{fv}(t)
\end{array}
$$

We use the above relations as examples throughout the paper. Of particular interest is the axiom $(\lambda x.E[x])\, t =_{\text{KH}} E[t]$ (if $x \notin \mathsf{fv}(E)$), called β_Ω in [13], which can be difficult to prove with bisimilarities [4].

2.4 Context Closures

Given a relation $\mathcal{R}$ on terms, we define two context closures that generate respectively terms and evaluation contexts. The term generating closure $\widehat{\mathcal{R}}$ is defined inductively as the smallest relation satisfying the following rules:

$$
\frac{t \mathrel{\mathcal{R}} t'}{t \mathrel{\widehat{\mathcal{R}}} t'} \qquad x \mathrel{\widehat{\mathcal{R}}} x \qquad \frac{t \mathrel{\widehat{\mathcal{R}}} t'}{\lambda x.t \mathrel{\widehat{\mathcal{R}}} \lambda x.t'} \qquad \frac{t_0 \mathrel{\widehat{\mathcal{R}}} t_0' \quad t_1 \mathrel{\widehat{\mathcal{R}}} t_1'}{t_0\, t_1 \mathrel{\widehat{\mathcal{R}}} t_0'\, t_1'} \qquad \frac{t \mathrel{\widehat{\mathcal{R}}} t'}{\mathcal{S}k.t \mathrel{\widehat{\mathcal{R}}} \mathcal{S}k.t'} \qquad \frac{t \mathrel{\widehat{\mathcal{R}}} t'}{\langle t \rangle \mathrel{\widehat{\mathcal{R}}} \langle t' \rangle}
$$

Even if $\mathcal{R}$ is defined only on closed terms, $\widehat{\mathcal{R}}$ is defined on open terms. In this paper, we consider the restriction of $\widehat{\mathcal{R}}$ to closed terms unless stated otherwise. The context generating closure $\widetilde{\mathcal{R}}$ of a relation $\mathcal{R}$ is defined inductively as the smallest relation satisfying the following rules:

$$
\frac{}{\Box \mathrel{\widetilde{\mathcal{R}}} \Box} \qquad \frac{F_0 \mathrel{\widetilde{\mathcal{R}}} F_1 \qquad v_0 \mathrel{\widehat{\mathcal{R}}} v_1}{v_0\ F_0 \mathrel{\widetilde{\mathcal{R}}} v_1\ F_1} \qquad \frac{F_0 \mathrel{\widetilde{\mathcal{R}}} F_1 \qquad t_0 \mathrel{\widehat{\mathcal{R}}} t_1}{F_0\ t_0 \mathrel{\widetilde{\mathcal{R}}} F_1\ t_1} \qquad \frac{F_0 \mathrel{\widetilde{\mathcal{R}}} F_1}{\langle F_0 \rangle \mathrel{\widetilde{\mathcal{R}}} \langle F_1 \rangle}
$$

Again, we consider only the restriction of $\widetilde{\mathcal{R}}$ to closed contexts.

3 Environmental Relations for the Relaxed Semantics

In this section, we define an environmental bisimilarity which characterizes the contextual equivalence of [4, 5], where stuck terms can be observed.

3.1 Contextual Equivalence

We recall the definition of contextual equivalence $\approx_c$ for the relaxed semantics (given in [4]).

Definition 5. *For all t_0, t_1 be terms. We write $t_0 \approx_c t_1$ if for all C such that $C[t_0]$ and $C[t_1]$ are closed, the following hold:*

- *$C[t_0] \Downarrow_{\mathrm{v}} v_0$ implies $C[t_1] \Downarrow_{\mathrm{v}} v_1$;*
- *$C[t_0] \Downarrow_{\mathrm{v}} t_0'$, where t_0' is stuck, implies $C[t_1] \Downarrow_{\mathrm{v}} t_1'$, with t_1' stuck as well;*

and conversely for $C[t_1]$.

The definition is simpler when using the following context lemma [15] (for a proof see Section 3.4 in [4]). Instead of testing with general, closing contexts, we can close the terms with values and then put them in evaluation contexts.

Lemma 3 (Context Lemma). *We have $t_0 \approx_c t_1$ iff for all closed contexts F and for all substitutions σ (mapping variables to closed values) such that $t_0\sigma$ and $t_1\sigma$ are closed, the following hold:*

- *$F[t_0\sigma] \Downarrow_{\mathrm{v}} v_0$ implies $F[t_1\sigma] \Downarrow_{\mathrm{v}} v_1$;*
- *$F[t_0\sigma] \Downarrow_{\mathrm{v}} t_0'$, where t_0' is stuck, implies $F[t_1\sigma] \Downarrow_{\mathrm{v}} t_1'$, with t_1' stuck as well;*

and conversely for $F[t_1\sigma]$.

In [4], we prove that $\approx_c$ satisfies all the axioms of CPS equivalence except for $\mathcal{S}k.k\,t =_{\mathrm{KH}} t$ (provided $k \notin \mathsf{fv}(t)$): indeed, $\mathcal{S}k.k\,t$ is stuck, but t may evaluate to a value. Conversely, some contextually equivalent terms are not CPS equivalent, like Turing's and Church's call-by-value fixed point combinators. Similarly, two arbitrary diverging terms are related by $\approx_c$, but not necessarily by $\equiv$.

3.2 Definition of Environmental Bisimulation and Basic Properties

Environmental bisimulations use an environment $\mathcal{E}$ to accumulate knowledge about two tested terms. For the λ-calculus [18], $\mathcal{E}$ records the values (v_0, v_1) the tested terms reduce to, if they exist. We can then compare v_0 and v_1 at any time by passing them arguments built from $\mathcal{E}$. In $\lambda_{\mathcal{S}}$, we have to consider stuck terms as well; therefore, environments may also contain pairs of stuck terms, and we can test those by building pure contexts from $\mathcal{E}$.

Formally, an environment $\mathcal{E}$ is a relation on normal forms which relates values with values and stuck terms with stuck terms; e.g., the identity environment $\mathcal{I}$ is $\{(t,t) \mid t \text{ is a normal form}\}$. An environmental relation $\mathcal{X}$ is a set of environments $\mathcal{E}$, and triples $(\mathcal{E}, t_0, t_1)$, where t_0 and t_1 are closed. We write $t_0 \; \mathcal{X}_{\mathcal{E}} \; t_1$ as a shorthand for $(\mathcal{E}, t_0, t_1) \in \mathcal{X}$; roughly, it means that we test t_0 and t_1 with the knowledge $\mathcal{E}$. The *open extension* of $\mathcal{X}$, written $\mathcal{X}^\circ$, is defined as follows: if $\overrightarrow{x} = \mathsf{fv}(t_0) \cup \mathsf{fv}(t_1)$[4], then we write $t_0 \; \mathcal{X}_{\mathcal{E}}^\circ \; t_1$ if $\lambda\overrightarrow{x}.t_0 \; \mathcal{X}_{\mathcal{E}} \; \lambda\overrightarrow{x}.t_1$.

[4] Given a metavariable m, we write $\overrightarrow{m}$ for a set of entities denoted by m.

Definition 6. *A relation $\mathcal{X}$ is an environmental bisimulation if*

1. $t_0 \; \mathcal{X}_{\mathcal{E}} \; t_1$ *implies:*
   (a) *if* $t_0 \rightarrow_{\mathrm{v}} t'_0$, *then* $t_1 \rightarrow^*_{\mathrm{v}} t'_1$ *and* $t'_0 \; \mathcal{X}_{\mathcal{E}} \; t'_1$;
   (b) *if* $t_0 = v_0$, *then* $t_1 \rightarrow^*_{\mathrm{v}} v_1$ *and* $\mathcal{E} \cup \{(v_0, v_1)\} \in \mathcal{X}$;
   (c) *if* t_0 *is stuck, then* $t_1 \rightarrow^*_{\mathrm{v}} t'_1$ *with* t'_1 *stuck, and* $\mathcal{E} \cup \{(t_0, t'_1)\} \in \mathcal{X}$;
   (d) *the converse of the above conditions on* t_1;
2. $\mathcal{E} \in \mathcal{X}$ *implies:*
   (a) *if* $\lambda x.t_0 \; \mathcal{E} \; \lambda x.t_1$ *and* $v_0 \; \widehat{\mathcal{E}} \; v_1$, *then* $t_0\{v_0/x\} \; \mathcal{X}_{\mathcal{E}} \; t_1\{v_1/x\}$;
   (b) *if* $E_0[\mathcal{S}k.t_0] \; \mathcal{E} \; E_1[\mathcal{S}k.t_1]$ *and* $E'_0 \; \widetilde{\mathcal{E}} \; E'_1$, *then* $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle \; \mathcal{X}_{\mathcal{E}}$ $\langle t_1\{\lambda x.\langle E'_1[E_1[x]]\rangle/k\}\rangle$ *for a fresh* x.

Environmental bisimilarity, written $\approx$, is the largest environmental bisimulation. To prove that two terms t_0 and t_1 are equivalent, we want to relate them without any predefined knowledge, i.e., we want to prove that $t_0 \approx_{\emptyset} t_1$ holds; we also write $\simeq$ for $\approx_{\emptyset}$.

The first part of the definition makes the bisimulation game explicit for t_0, t_1, while the second part focuses on environments $\mathcal{E}$. If t_0 is a normal form, then t_1 has to evaluate to a normal form of the same kind, and we extend the environment with the newly acquired knowledge. We then compare values in $\mathcal{E}$ (clause (2a)) by applying them to arguments built from $\mathcal{E}$, as in the λ-calculus [18]. Similarly, we test stuck terms in $\mathcal{E}$ by putting them within contexts $\langle E'_0\rangle$, $\langle E'_1\rangle$ built from $\mathcal{E}$ (clause (2b)) to trigger the capture. This reminds the way we test values and stuck terms with applicative bisimilarity [4], except that applicative bisimilarity tests both values or stuck terms with the same argument or context. Using different entities (as in Definition 6) makes bisimulation proofs harder, but it simplifies the proof of congruence of the environmental bisimilarity.

Example 1. We have $\langle(\lambda x.t_0)\,\langle t_1\rangle\rangle \simeq (\lambda x.\langle t_0\rangle)\,\langle t_1\rangle$, because the relation $\mathcal{X} = \{(\emptyset, \langle(\lambda x.t)\,\langle t'\rangle\rangle, (\lambda x.\langle t\rangle)\,\langle t'\rangle), (\emptyset, \langle(\lambda x.t)\,v\rangle, (\lambda x.\langle t\rangle) v)\} \cup \{(\mathcal{E}, t, t) \mid \mathcal{E} \subseteq \mathcal{I}\} \cup \{\mathcal{E} \mid \mathcal{E} \subseteq \mathcal{I}\}$ is a bisimulation. Indeed, if $\langle t'\rangle$ evaluates to v, then $\langle(\lambda x.t)\,\langle t'\rangle\rangle \rightarrow^*_{\mathrm{v}} \langle(\lambda x.t)\,v\rangle$ and $(\lambda x.\langle t\rangle)\,\langle t'\rangle \rightarrow^*_{\mathrm{v}} (\lambda x.\langle t\rangle)\,v$, which both reduce to $\langle t\{v/x\}\rangle$.

As usual with environmental relations, the candidate relation $\mathcal{X}$ in the above example could be made simpler with the help of up-to techniques.

Definition 6 is written in the small-step style, because each reduction step from t_0 has to be matched by t_1. In the big-step style, we are concerned only with evaluations to normal forms.

Definition 7. *A relation $\mathcal{X}$ is a big-step environmental bisimulation if* $t_0 \; \mathcal{X}_{\mathcal{E}} \; t_1$ *implies:*

1. $t_0 \; \mathcal{X}_{\mathcal{E}} \; t_1$ *implies:*
   (a) *if* $t_0 \rightarrow^*_{\mathrm{v}} v_0$, *then* $t_1 \rightarrow^*_{\mathrm{v}} v_1$ *and* $\mathcal{E} \cup \{(v_0, v_1)\} \in \mathcal{X}$;
   (b) *if* $t_0 \rightarrow^*_{\mathrm{v}} t'_0$ *with* t'_0 *stuck, then* $t_1 \rightarrow^*_{\mathrm{v}} t'_1$, t'_1 *stuck, and* $\mathcal{E} \cup \{(t'_0, t'_1)\} \in \mathcal{X}$;
   (c) *the converse of the above conditions on* t_1;
2. $\mathcal{E} \in \mathcal{X}$ *implies:*
   (a) *if* $\lambda x.t_0 \; \mathcal{E} \; \lambda x.t_1$ *and* $v_0 \; \widehat{\mathcal{E}} \; v_1$, *then* $t_0\{v_0/x\} \; \mathcal{X}_{\mathcal{E}} \; t_1\{v_1/x\}$;

(b) if $E_0[\mathcal{S}k.t_0]\ \mathcal{E}\ E_1[\mathcal{S}k.t_1]$ *and* $E'_0\ \widetilde{\mathcal{E}}\ E'_1$*, then* $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle\ \mathcal{X}_{\mathcal{E}}$ $\langle t_1\{\lambda x.\langle E'_1[E_1[x]]\rangle/k\}\rangle$ *for a fresh* x*.*

Lemma 4. *If* $\mathcal{X}$ *is a big-step environmental bisimulation, then* $\mathcal{X} \subseteq \approx$*.*

Big-step relations can be more convenient to use when we know the result of the evaluation, as in Example 1, or as in the following one.

Example 2. We have $\langle\langle t\rangle\rangle \simeq \langle t\rangle$. Indeed, we can show that $\langle\langle t\rangle\rangle \rightarrow^*_{\mathrm{v}} v$ iff $\langle t\rangle \rightarrow^*_{\mathrm{v}} v$, therefore $\{(\emptyset, \langle\langle t\rangle\rangle, \langle t\rangle)\} \cup \{(\mathcal{E}, t, t) \mid \mathcal{E} \subseteq \mathcal{I}\} \cup \{\mathcal{E} \mid \mathcal{E} \subseteq \mathcal{I}\}$ is a big-step environmental bisimulation.

We use the following results in the rest of the paper.

Lemma 5 (Weakening). *If* $t_0 \approx_{\mathcal{E}} t_1$ *and* $\mathcal{E}' \subseteq \mathcal{E}$ *then* $t_0 \approx_{\mathcal{E}'} t_1$*.*

A smaller environment is a weaker constraint, because we can build less arguments and contexts to test the normal forms in $\mathcal{E}$. The proof is as in [18]. Lemma 6 states that reduction (and therefore, evaluation) is included in $\simeq$.

Lemma 6. *If* $t_0 \rightarrow_{\mathrm{v}} t'_0$*, then* $t_0 \simeq t'_0$*.*

3.3 Soundness and Completeness

We now prove soundness and completeness of $\simeq$ w.r.t. contextual equivalence. Because the proofs follow the same steps as for the λ-calculus [18], we only give here the main lemmas and sketch their proofs. The complete proofs can be found in Appendix A. First, we need some basic up-to techniques, namely up-to environment (which allows bigger environments in the bisimulation clauses) and up-to bisimilarity (which allows for limited uses of $\simeq$ in the bisimulation clauses), whose definitions and proofs of soundness are classic [18].

With these tools, we can prove that $\simeq$ is sound and complete w.r.t. contextual equivalence. For a relation $\mathcal{R}$ on terms, we write $\mathcal{R}^{\mathrm{nf}}$ for its restriction to closed normal forms. The first step consists in proving congruence for normal forms, and also for any terms but only w.r.t. evaluation contexts.

Lemma 7. *Let* t_0*,* t_1 *be normal forms. If* $t_0 \approx_{\mathcal{E}} t_1$*, then* $C[t_0] \approx_{\mathcal{E}} C[t_1]$*.*

Lemma 8. *If* $t_0 \approx_{\mathcal{E}} t_1$*, then* $F[t_0] \approx_{\mathcal{E}} F[t_1]$*.*

Lemmas 7 and 8 are proved simultaneously by showing that, for any environmental bisimulation $\mathcal{Y}$, the relation

$$\mathcal{X} = \{(\widehat{\mathcal{E}}^{\mathrm{nf}}, F_0[t_0], F_1[t_1]) \mid t_0\ \mathcal{Y}_{\mathcal{E}}\ t_1, F_0\ \widetilde{\mathcal{E}}\ F_1\} \\ \cup \{(\widehat{\mathcal{E}}^{\mathrm{nf}}, t_0, t_1) \mid \mathcal{E} \in \mathcal{Y}, t_0\ \widehat{\mathcal{E}}\ t_1\} \cup \{\widehat{\mathcal{E}}^{\mathrm{nf}} \mid \mathcal{E} \in \mathcal{Y}\}$$

is a bisimulation up-to environment. Informally, the elements of the first set of $\mathcal{X}$ reduce to elements of the second set of $\mathcal{X}$, and we then prove the bisimulation property for these elements by induction on $t_0\ \widehat{\mathcal{E}}\ t_1$. We can then prove the main congruence lemma.

Lemma 9. $t_0 \simeq t_1$ *implies* $C[t_0] \approx_{\widehat{\simeq}^{\text{nf}}} C[t_1]$.

We show that $\{(\widehat{\simeq}^{\text{nf}}, t_0, t_1) \mid t_0 \widehat{\simeq} t_1\} \cup \{\widehat{\simeq}^{\text{nf}}\}$ is a bisimulation up-to bisimilarity by induction on $t_0 \widehat{\simeq} t_1$. By weakening (Lemma 5), we can deduce from Lemma 9 that $\simeq$ is a congruence, and therefore is sound w.r.t. $\approx_c$.

Corollary 1 (Soundness). *We have* $\simeq \subseteq \approx_c$.

The relation $\simeq$ is also complete w.r.t. contextual equivalence.

Theorem 1 (Completeness). *We have* $\approx_c \subseteq \simeq$.

The proof is by showing that $\{(\approx_c{}^{\text{nf}}, t_0, t_1) \mid t_0 \approx_c t_1\} \cup \{\approx_c{}^{\text{nf}}\}$ is a big-step bisimulation, using Lemma 3 as an alternate definition for $\approx_c$.

3.4 Bisimulation up to context

Equivalence proofs based on environmental bisimilarity can be simplified by using up-to techniques, such as up to reduction, up to expansion, and up to context [18]. We only discuss the last, since the first two can be defined and proved sound in $\lambda_{\mathcal{S}}$ without issues. Bisimulations up to context may factor out a common context from the tested terms. Formally, we define the context closure of $\mathcal{X}$, written $\overline{\mathcal{X}}$, as follows: we have $t_0 \; \overline{\mathcal{X}_{\mathcal{E}}} \; t_1$ if

- either $t_0 = F_0[t'_0]$, $t_1 = F_1[t'_1]$, $t'_0 \; \mathcal{X}_{\mathcal{E}} \; t'_1$, and $F_0 \; \widetilde{\mathcal{E}} \; F_1$;
- or $t_0 \; \widehat{\mathcal{E}} \; t_1$.

Note that terms t'_0 and t'_1 (related by $\mathcal{X}_{\mathcal{E}}$) can be put into evaluation contexts only, while normal forms (related by $\mathcal{E}$) can be put in any contexts. This restriction to evaluation contexts in the first case is usual in the definition of up-to context techniques for environmental relations [18, 21, 19, 17].

Definition 8. *A relation* $\mathcal{X}$ *is an environmental bisimulation up to context if*

1. $t_0 \; \mathcal{X}_{\mathcal{E}} \; t_1$ *implies:*
   (a) *if* $t_0 \rightarrow_{\text{v}} t'_0$, *then* $t_1 \rightarrow^*_{\text{v}} t'_1$ *and* $t'_0 \; \overline{\mathcal{X}_{\mathcal{E}}} \; t'_1$;
   (b) *if* $t_0 = v_0$, *then* $t_1 \rightarrow^*_{\text{v}} v_1$ *and* $\mathcal{E} \cup \{(v_0, v_1)\} \subseteq \widehat{\mathcal{E}'}^{\text{nf}}$ *for some* $\mathcal{E}' \in \mathcal{X}$;
   (c) *if* t_0 *is stuck, then* $t_1 \rightarrow^*_{\text{v}} t'_1$ *with* t'_1 *stuck, and* $\mathcal{E} \cup \{(t_0, t'_1)\} \subseteq \widehat{\mathcal{E}'}^{\text{nf}}$ *for some* $\mathcal{E}' \in \mathcal{X}$;
   (d) *the converse of the above conditions on* t_1;
2. $\mathcal{E} \in \mathcal{X}$ *implies:*
   (a) *if* $\lambda x.t_0 \; \mathcal{E} \; \lambda x.t_1$ *and* $v_0 \; \widehat{\mathcal{E}} \; v_1$, *then* $t_0\{v_0/x\} \; \overline{\mathcal{X}_{\mathcal{E}}} \; t_1\{v_1/x\}$;
   (b) *if* $E_0[\mathcal{S}k.t_0] \; \mathcal{E} \; E_1[\mathcal{S}k.t_1]$ *and* $E'_0 \; \widetilde{\mathcal{E}} \; E'_1$, *then* $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle \; \overline{\mathcal{X}_{\mathcal{E}}} \; \langle t_1\{\lambda x.\langle E'_1[E_1[x]]\rangle/k\}\rangle$ *for a fresh* x.

Lemma 10. *If* $\mathcal{X}$ *is an environmental bisimulation up to context, then* $\mathcal{X} \subseteq \approx$.

The soundness proof is the same as in [18]. While this definition is enough to simplify proofs in the λ-calculus case, it is not that helpful in $\lambda_{\mathcal{S}}$, because of the restriction to evaluation contexts (first item of the definition of $\overline{\mathcal{X}}$). In the λ-calculus, when a term t reduces within an evaluation context, the context is not affected, hence Definition 8 is enough to help proving interesting equivalences. It is not the case in $\lambda_{\mathcal{S}}$, as (a part of) the evaluation context can be captured.

Indeed, suppose we want to construct a candidate relation $\mathcal{X}$ to prove the β_{Ω} axiom, i.e., $E[t]$ is equivalent to $(\lambda x.E[x])\ t$, assuming $x \notin \mathsf{fv}(E)$. The problematic case is when t is a stuck term $E_0[\mathcal{S}k.t_0]$; we have to add the stuck terms $(\lambda x.E[x])\ E_0[\mathcal{S}k.t_0]$ and $E[E_0[\mathcal{S}k.t_0]]$ to an environment $\mathcal{E}$ of $\mathcal{X}$. For $\mathcal{X}$ to be a bisimulation, we then have to prove that for all $E_1\ \widetilde{\mathcal{E}}\ E_2$, we have $\langle t_0\{\lambda y.\langle E_1[(\lambda x.E[x])\ E_0[y]]\rangle/k\}\rangle\ \mathcal{X}_{\mathcal{E}}\ \langle t_0\{\lambda y.\langle E_2[E[E_0[y]]]\rangle/k\}\rangle$. At this point, we would like to use the up-to context technique, because the subterms $(\lambda x.E[x])\,E_0[y]$ and $E[E_0[y]]$ are similar to the terms we want to relate (they can be written $(\lambda x.E[x])\ t''$ and $E[t'']$ with $t'' = E_0[y]$). However, we have at best $\langle t_0\{\lambda y.\langle E_1[(\lambda x.E[x])\ E_0[y]]\rangle/k\}\rangle\ \widehat{\mathcal{X}_{\mathcal{E}}}^{\circ}\ \langle t_0\{\lambda y.\langle E_2[E[E_0[y]]]\rangle/k\}\rangle$ (and not $\overline{\mathcal{X}_{\mathcal{E}}}$), because (i) $(\lambda x.E[x])\ E_0[y]$ and $E[E_0[y]]$ are open terms, and (ii) t_0 can be any term, so $(\lambda x.E[x])\,E_0[y]$ and $E[E_0[y]]$ can be put in any context, not necessarily in an evaluation one. Therefore, Definition 8 cannot help there.

Problem (ii) could be somewhat dealt with in the particular case of the β_{Ω} axiom by changing clause (2b) of Definition 8 into

(b) if $E_0[\mathcal{S}k.t_0]\ \mathcal{E}\ E_1[\mathcal{S}k.t_1]$ and $E'_0\ \widetilde{\mathcal{X}_{\mathcal{E}}}\ E'_1$, then $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle\ \widehat{\mathcal{X}_{\mathcal{E}}}$ $\langle t_1\{\lambda x.\langle E'_1[E_1[x]]\rangle/k\}\rangle$ for a fresh x.

and similarly for clause (2a). In plain text, we build the testing contexts E'_0, E'_1 from $\mathcal{X}_{\mathcal{E}}$ (instead of $\mathcal{E}$), and the resulting terms have to be in $\widehat{\mathcal{X}_{\mathcal{E}}}$ (without any evaluation context restriction). The resulting notion of bisimulation up to context is sound. The new clause would be more difficult to establish in general than the original one (of Definition 8), because it tests more pairs of contexts. However, for the β_{Ω} axiom, we would have to prove that for all $E_1\ \widehat{\mathcal{X}_{\mathcal{E}}}\ E_2$, $\langle t_0\{\lambda y.\langle E_1[(\lambda x.E[x])\ E_0[y]]\rangle/k\}\rangle\ \widehat{\mathcal{X}_{\mathcal{E}}}\ \langle t_0\{\lambda y.\langle E_2[E[E_0[y]]]\rangle/k\}\rangle$ holds; it would be easy, except $(\lambda x.E[x])\ E_0[y]$ and $E[E_0[y]]$ are open terms (problem (i)).

Problem (i) seems harder to fix, because for $(\lambda x.E[x])\ E_0[y]\ \mathcal{X}_{\mathcal{E}}{}^{\circ}\ E[E_0[y]]$ to hold, we must have $(\lambda x.E[x])\ E_0[v_0]\ \mathcal{X}_{\mathcal{E}}\ E[E_0[v_1]]$ for all $v_0\ \widehat{\mathcal{E}}\ v_1$. Because E_0 can be anything, it means that we must have $(\lambda x.E[x])\,t'_0\ \mathcal{X}_{\mathcal{E}}\ E[t'_1]$ with $t'_0\ \widehat{\mathcal{E}}\ t'_1$; t'_0 and t'_1 are plugged in different contexts, therefore bisimulation up to context (which factors out only a common context) cannot help us there; a new kind of up-to technique is required.

The β_{Ω} axiom example suggests that we need more powerful up-to techniques for environmental bisimilarity for delimited control; we leave these potential improvements as a future work. Note that we do not have such issues with up-to techniques for normal form bisimilarity: it relates open terms without having to replace their free variables, and normal form bisimulation up to context is not restricted to evaluation contexts only. But even if environmental bisimulation up to context is not as helpful as wished, it still simplifies equivalence proofs, as we can see with the next example.

Example 3. In [6], a variant of Turing's call-by-value fixed point combinators using shift and reset has been proposed. Let $\theta = \lambda xy.y\ (\lambda z.x\ x\ y\ z)$. We prove that $t_0 = \theta\ \theta$ is bisimilar to its variant $t_1 = \langle \theta\ \mathcal{S}k.k\ k\rangle$. Let $\theta' = \lambda x.\langle \theta\ x\rangle$, $v_0 = \lambda y.y\ (\lambda z.\theta\ \theta\ y\ z)$, and $v_1 = \lambda y.y\ (\lambda z.\theta'\ \theta'\ y\ z)$. We define $\mathcal{E}$ inductively such that $v_0\ \mathcal{E}\ v_1$, and if $v_0'\ \widehat{\mathcal{E}}\ v_1'$, then $\lambda z.\theta\ \theta\ v_0'\ z\ \mathcal{E}\ \lambda z.\theta'\ \theta'\ v_1'\ z$. Then $\mathcal{X} = \{(\mathcal{E}, t_0, t_1), (\mathcal{E}, t_0, \theta'\ \theta'), \mathcal{E}\}$ is a (big-step) bisimulation up to context. Indeed, we have $t_0 \Downarrow_{\mathrm{v}} v_0$, $t_1 \Downarrow_{\mathrm{v}} v_1$, and $\theta'\ \theta' \Downarrow_{\mathrm{v}} v_1$, therefore clause (1b) of Definition 8 is checked for both pairs. We now check clause (2a), first for $v_0\ \mathcal{E}\ v_1$. For all $v_0'\ \widehat{\mathcal{E}}\ v_1'$, we have $v_0'\ (\lambda z.\theta\ \theta\ v_0'\ z)\ \widehat{\mathcal{E}}\ v_1'\ (\lambda z.\theta'\ \theta'\ v_1'\ z)$ (because $\lambda z.\theta\ \theta\ v_0'\ z\ \mathcal{E}\ \lambda z.\theta'\ \theta'\ v_1'\ z$), hence the result holds. Next, let $\lambda z.\theta\,\theta\, v_0'\, z\, \mathcal{E}\, \lambda z.\theta'\,\theta'\, v_1'\, z$ (with $v_0'\ \widehat{\mathcal{E}}\ v_1'$), and let $v_0''\ \widehat{\mathcal{E}}\ v_1''$. We have to check that $\theta\ \theta\ v_0'\ v_0''\ \overline{\mathcal{X}_{\mathcal{E}}}\ \theta'\ \theta'\ v_1'\ v_1''$, which is true, because $\theta\ \theta\ \mathcal{X}_{\mathcal{E}}\ \theta'\ \theta'$, and $\Box\ v_0'\ v_0''\ \widetilde{\mathcal{E}}\ \Box\ v_1'\ v_1''$.

4 Environmental Relations for the Original Semantics

The original CPS semantics for shift and reset [7] as well as the corresponding reduction semantics [3] assume that terms can be considered as programs to be executed, only when surrounded by a top-level reset. In this section, we present a CPS-compatible bisimulation theory that takes such a requirement into account. In this section, we call *programs*, ranged over by p, terms of the form $\langle t\rangle$.

4.1 Contextual Equivalence

To reflect the fact that terms are executed within an enclosing reset, the contextual equivalence we consider in this section tests terms in contexts of the form $\langle C\rangle$ only. Because programs cannot reduce to stuck terms, the only possible observable action is evaluation to values. We therefore define contextual equivalence for programs as follows.

Definition 9. *Let t_0, t_1 be terms. We write $t_0 \dot{\approx}_c t_1$ if for all C such that $\langle C[t_0]\rangle$ and $\langle C[t_1]\rangle$ are closed, $\langle C[t_0]\rangle \Downarrow_{\mathrm{v}} v_0$ implies $\langle C[t_1]\rangle \Downarrow_{\mathrm{v}} v_1$, and conversely for $\langle C[t_1]\rangle$.*

Note that $\dot{\approx}_c$ is defined on all terms, not just programs. It is easy to check that $\approx_c$ is more discriminative than $\dot{\approx}_c$. We will see in Section 4.4 that this inclusion is in fact strict.

Lemma 11. *We have $\approx_c \subseteq \dot{\approx}_c$.*

4.2 Definition and Properties

We now propose a definition of environmental bisimulation adapted to programs (but defined on all terms, like $\dot{\approx}_c$). Because stuck terms are no longer observed, environments $\mathcal{E}$ henceforth relate only values. Similarly, we write $\mathcal{R}^{\mathrm{v}}$ for the restriction of a relation $\mathcal{R}$ on terms to pairs of closed values.

Definition 10. *A relation $\mathcal{X}$ is an environmental bisimulation for programs if*

1. *if $t_0 \; \mathcal{X}_\mathcal{E} \; t_1$ and t_0 and t_1 are not both programs, then for all $E_0 \; \widetilde{\mathcal{E}} \; E_1$, we have $\langle E_0[t_0]\rangle \; \mathcal{X}_\mathcal{E} \; \langle E_1[t_1]\rangle$;*
2. *if $p_0 \; \mathcal{X}_\mathcal{E} \; p_1$*
   (a) if $p_0 \rightarrow_\mathrm{v} p_0'$, then $p_1 \rightarrow_\mathrm{v}^* p_1'$ and $p_0' \; \mathcal{X}_\mathcal{E} \; p_1'$;
   (b) if $p_0 \rightarrow_\mathrm{v} v_0$, then $p_1 \rightarrow_\mathrm{v}^* v_1$, and $\{(v_0, v_1)\} \cup \mathcal{E} \in \mathcal{X}$;
   (c) the converse of the above conditions on p_1;
3. *for all $\mathcal{E} \in \mathcal{X}$, if $\lambda x.t_0 \; \mathcal{E} \; \lambda x.t_1$ and $v_0 \; \widehat{\mathcal{E}} \; v_1$, then $t_0\{v_0/x\} \; \mathcal{X}_\mathcal{E} \; t_1\{v_1/x\}$.*

Environmental bisimilarity for programs, written $\dot{\approx}$, is the largest environmental bisimulation for programs. As before, the relation $\dot{\approx}_\emptyset$, also written $\dot{\simeq}$, is candidate to characterize $\dot{\approx}_c$.

Clauses (2) and (3) of Definition 10 deal with programs and environment in a classical way (as in plain λ-calculus). The problematic case is when relating terms t_0 and t_1 that are not both programs (clause (1)). Indeed, one of them may be stuck, and therefore we have to test them within some contexts $\langle E_0\rangle$, $\langle E_1\rangle$ (built from $\mathcal{E}$) to potentially trigger a capture that otherwise would not happen. We cannot require both terms to be stuck, as in clause (2b) of Definition 6, because a stuck term can be equivalent to a term free from control effect. E.g., we will see that $v \mathrel{\dot{\simeq}} \mathcal{S}k.k\,v$, provided that $k \notin \mathsf{fv}(v)$.

Example 4. Suppose we want to prove $\langle(\lambda x.t_0)\,\langle t_1\rangle\rangle \mathrel{\dot{\simeq}} (\lambda x.\langle t_0\rangle)\,\langle t_1\rangle$ (as in Example 1). Because $(\lambda x.\langle t_0\rangle)\,\langle t_1\rangle$ is not a program, we have to put both terms into a context first: we have to change the candidate relation of Example 1 into $\mathcal{X} = \{(\emptyset, \langle(\lambda x.t_0)\,\langle t_1\rangle\rangle, (\lambda x.\langle t_0\rangle)\,\langle t_1\rangle)\} \cup \{(\emptyset, \langle E[\langle(\lambda x.t_0)\,\langle t_1\rangle\rangle]\rangle, \langle E[(\lambda x.\langle t_0\rangle)\,\langle t_1\rangle]\rangle)\} \cup \{(\emptyset, \langle E[\langle(\lambda x.t_0)\,v\rangle]\rangle, \langle E[(\lambda x.\langle t_0\rangle)\,v]\rangle)\} \cup \{(\mathcal{E}, t, t) \mid \mathcal{E} \subseteq \mathcal{I}\} \cup \{\mathcal{E} \mid \mathcal{E} \subseteq \mathcal{I}\}$. In contrast, to prove $\langle\langle t\rangle\rangle \mathrel{\dot{\simeq}} \langle t\rangle$, we do not have to change the candidate relation of Example 2, since both terms are programs.

We can give a definition of big-step bisimulation by removing clause (2a) and changing $\rightarrow_\mathrm{v}$ into $\rightarrow_\mathrm{v}^*$ in clause (2b). Lemmas 5 and 6 can also be extended to $\dot{\approx}$ and $\dot{\simeq}$. The next lemma shows that $\simeq$ is more discriminative than $\dot{\simeq}$.

Lemma 12. *We have $\simeq \subseteq \dot{\simeq}$.*

A consequence of Lemma 12 is that we can use Definition 6 as a proof technique for $\dot{\simeq}$. E.g., we have directly $\langle(\lambda x.t_0)\,\langle t_1\rangle\rangle \mathrel{\dot{\simeq}} (\lambda x.\langle t_0\rangle)\,\langle t_1\rangle$, because $\langle(\lambda x.t_0)\,\langle t_1\rangle\rangle \simeq (\lambda x.\langle t_0\rangle)\,\langle t_1\rangle$.

4.3 Soundness and Completeness

We sketch the proofs of soundness and completeness of $\dot{\simeq}$ w.r.t. $\dot{\approx}_c$; see Appendix B for the complete proofs. The soundness proof follows the same scheme as in Section 3.3, with some necessary adjustments. As before, we need up-to environment and up-to bisimilarity techniques to prove the following lemmas.

Lemma 13. *If $v_0 \mathrel{\dot{\approx}_\mathcal{E}} v_1$, then $C[v_0] \mathrel{\dot{\approx}_\mathcal{E}} C[v_1]$.*

Lemma 14. *If $t_0 \dot{\approx}_{\mathcal{E}} t_1$, then $F[t_0] \dot{\approx}_{\mathcal{E}} F[t_1]$.*

We prove Lemmas 13 and 14 by showing that a relation similar to the relation $\mathcal{X}$ defined in Section 3.3 is a bisimulation up to environment. We then want to prove the main congruence lemma, akin to Lemma 9, by showing that $\mathcal{Y} = \{(\widehat{\dot{\simeq}}^{\mathrm{v}}, t_0, t_1) \mid t_0 \mathrel{\widehat{\dot{\simeq}}} t_1\} \cup \{\widehat{\dot{\simeq}}^{\mathrm{v}}\}$ is a bisimulation up to bisimilarity. However, we can no longer proceed by induction on $t_0 \mathrel{\widehat{\dot{\simeq}}} t_1$, as for Lemma 9. Indeed, if $p_0 = \langle t_0 \rangle$, $p_1 = \langle t_1 \rangle$ with $t_0 \mathrel{\widehat{\dot{\simeq}}} t_1$, and if t_0 is a stuck term, then p_0 reduces to some term, but the induction hypothesis does not tell us anything about t_1. To circumvent this, we decompose related programs into related subcomponents.

Lemma 15. *If $p_0 \mathrel{\widehat{\dot{\simeq}}} p_1$, then either $p_0 \dot{\simeq} p_1$, or one of the following holds:*

- $p_0 = \langle v_0 \rangle$;
- $p_0 = F_0[\langle E_0[t_0] \rangle]$, $p_1 = F_1[\langle E_1[t_1] \rangle]$, $F_0 \mathrel{\widetilde{\dot{\simeq}}} F_1$, $E_0 \mathrel{\widetilde{\dot{\simeq}}} E_1$, $t_0 \dot{\simeq} t_1$ *and* $t_0 \rightarrow_{\mathrm{v}} t_0'$ *or* t_0 *is stuck;*
- $p_0 = F_0[\langle E_0[r_0] \rangle]$, $p_1 = F_1[\langle E_1[t_1] \rangle]$, $F_0 \mathrel{\widetilde{\dot{\simeq}}} F_1$, $E_0 \mathrel{\widetilde{\dot{\simeq}}} E_1$, $r_0 \mathrel{\widehat{\dot{\simeq}}} t_1$ *but* $r_0 \not\dot{\simeq} t_1$.

Lemma 15 generalizes Lemma 2 to related programs: we know p_0 can be decomposed into contexts F, $\langle E \rangle$, and a redex r, and we relate these subterms to p_1. We can then prove that $\mathcal{Y}$ (defined above) is a bisimulation up to bisimilarity, by showing that, in each case described by Lemma 15, p_0 and p_1 reduce to terms related by $\mathcal{Y}$. From this, we deduce $\dot{\simeq}$ is a congruence, and is sound w.r.t. $\dot{\approx}_c$.

Lemma 16. *$t_0 \dot{\simeq} t_1$ implies $C[t_0] \dot{\approx}_{\widehat{\dot{\simeq}}^{\mathrm{v}}} C[t_1]$.*

Corollary 2 (Soundness). *We have $\dot{\simeq} \subseteq \dot{\approx}_c$.*

Remark 1. Following the ideas behind Definition 10, one can define an applicative bisimilarity $\mathcal{B}$ for programs. However, proving that $\mathcal{B}$ is sound seems more complex than for $\dot{\simeq}$. We remind that the soundness proof of an applicative bisimilarity consists in showing that a relation called the *Howe's closure* $\mathcal{B}^\bullet$ is an applicative bisimulation. To this end, we need a version of Lemma 15 for $\mathcal{B}^\bullet$. However, $\mathcal{B}^\bullet$ is inductively defined as the smallest congruence which contains $\mathcal{B}$ and satisfies $\mathcal{B}^\bullet\mathcal{B} \subseteq \mathcal{B}^\bullet$ (1), and condition (1) makes it difficult to write a decomposition lemma for $\mathcal{B}^\bullet$ similar to Lemma 15.

We prove completeness of $\dot{\simeq}$ by showing that the relation $\ddot{\approx}_c$, defined below, coincides with $\dot{\approx}_c$ and $\dot{\simeq}$. By doing so, we also prove a context lemma for $\dot{\approx}_c$.

Definition 11. *Let t_0, t_1 be closed terms. We write $t_0 \ddot{\approx}_c t_1$ if for all closed F, $\langle F[t_0] \rangle \Downarrow_{\mathrm{v}} v_0$ implies $\langle F[t_1] \rangle \Downarrow_{\mathrm{v}} v_1$, and conversely for $\langle F[t_1] \rangle$.*

By definition, we have $\dot{\approx}_c \subseteq \ddot{\approx}_c$. With the same proof technique as in Section 3.3, we prove the following lemma.

Lemma 17 (Completeness). *We have $\ddot{\approx}_c \subseteq \dot{\simeq}$.*

With Lemma 17 and Corollary 2, we have $\dot{\approx}_c \subseteq \ddot{\approx}_c \subseteq \dot{\simeq} \subseteq \dot{\approx}_c$. Defining up-to context for programs is possible, with the same limitations as in Section 3.4.

4.4 Examples

We illustrate the differences between $\simeq$ and $\dot{\simeq}$, by giving some examples of terms related by $\dot{\simeq}$, but not by $\simeq$. First, note that $\dot{\simeq}$ relates non-terminating terms with stuck non-terminating terms.

Lemma 18. *We have $\Omega \dot{\simeq} \mathcal{S}k.\Omega$.*

The relation $\{(\emptyset, \Omega, \mathcal{S}k.\Omega), (\emptyset, \langle E[\Omega]\rangle, \langle E[\mathcal{S}k.\Omega]\rangle), (\emptyset, \langle E[\Omega]\rangle, \langle \Omega\rangle)\}$ is a bisimulation for programs. Lemma 18 does not hold with $\simeq$ because Ω is not stuck.

As wished, $\dot{\simeq}$ satisfies the only axiom of [13] not satisfied by $\simeq$.

Lemma 19. *If $k \notin \mathsf{fv}(t)$, then $t \dot{\simeq}^{\circ} \mathcal{S}k.k\,t$.*

We sketch the proof for t closed; for the general case, see Appendix C.1. We prove that $\{(\emptyset, t, \mathcal{S}k.k\,t), (\emptyset, \langle E[t]\rangle, \langle E[\mathcal{S}k.k\,t]\rangle)\} \cup \simeq$ is a bisimulation for programs. Indeed, we have $\langle E[\mathcal{S}k.k\,t]\rangle \rightarrow_{\mathrm{v}} \langle(\lambda x.\langle E[x]\rangle)\,t\rangle$, and because $\simeq$ verifies the β_Ω axiom ($\simeq$ is complete, and $\approx_c$ verifies the β_Ω axiom [4]), we know that $\langle(\lambda x.\langle E[x]\rangle)\,t\rangle \simeq \langle\langle E[t]\rangle\rangle$ holds. From Example 2, we have $\langle\langle E[t]\rangle\rangle \simeq \langle E[t]\rangle$, therefore we have $\langle E[\mathcal{S}k.k\,t]\rangle \simeq \langle E[t]\rangle$.

Consequently, $\dot{\simeq}^{\circ}$ is complete w.r.t. $\equiv$.

Corollary 3. *We have $\equiv \subseteq \dot{\simeq}^{\circ}$.*

As a result, we can use $\equiv$ (restricted to closed terms) as a proof technique for $\dot{\simeq}$. E.g., the following equivalence can be derived from the axioms [13].

Lemma 20. *If $k \notin \mathsf{fv}(t_1)$, then $(\lambda x.\mathcal{S}k.t_0)\,t_1 \dot{\simeq} \mathcal{S}k.((\lambda x.t_0)\,t_1)$.*

This equivalence does not hold with $\simeq$, because the term on the right is stuck, but the term on the left may not evaluate to a stuck term (if t_1 does not terminate). We can generalize this result as follows, again by using $\equiv$.

Lemma 21. *If $k \notin \mathsf{fv}(t_1)$ and $x \notin \mathsf{fv}(E)$, then we have $(\lambda x.E[\mathcal{S}k.t_0])\,t_1 \dot{\simeq} E[\mathcal{S}k.((\lambda x.t_0)\,t_1)]$.*

Proving Lemma 19 without the β_Ω axiom and Lemmas 20 and 21 without $\equiv$ requires complex candidate relations (see the proof of Lemma 20 in Appendix C.2), because of the lack of powerful enough up-to techniques.

5 Conclusion

We propose sound and complete environmental bisimilarities for two variants of the semantics of $\lambda_{\mathcal{S}}$. For the semantics of Section 3, we now have several bisimilarities, each with its own merit. Normal form bisimilarity [5] and its up-to techniques leads to minimal proof obligations, however it is not complete, and distinguishes very simple equivalent terms (see Proposition 1 in [5]). Applicative bisimilarity [4] is complete but sometimes requires complex bisimulation proofs (e.g., for the β_Ω axiom). Environmental bisimilarity $\simeq$ (Definition 6) is also complete, can be difficult to use, but this difficulty can be mitigated with up-to techniques. However, bisimulation up to context is not as helpful as we could hope (see Section 3.4), because we have to manipulate open terms (problem (i)), and the context closure of an environmental relation is restricted to evaluation contexts (problem (ii)). As a result, proving the β_Ω axiom is more difficult with environmental than with applicative bisimilarity. We believe dealing with problem (i) requires new up-to techniques to be developed, and lifting the evaluation context restriction (problem (ii)) would benefit not only for $\lambda_{\mathcal{S}}$, but also for process calculi with passivation [17]; we leave this as a future work.

In contrast, we do not have as many options when considering the semantics of Section 4 (where terms are evaluated within a top-level reset). The environmental bisimilarity of this paper $\dot{\simeq}$ (Definition 10) is the first to be sound and complete w.r.t. Definition 9. As argued in [5] (Section 3.2), normal form bisimilarity cannot be defined on programs without introducing extra quantifications (which defeats the purpose of normal form bisimilarity). Applicative bisimilarity could be defined for programs, but proving its soundness would require a new technique, since the usual one (Howe's method) does not seem to apply (see Remark 1). This confirms that environmental bisimilarity is more flexible than applicative bisimilarity [14]. However, we would like to simplify the quantification over contexts in clause (1) of Definition 10, so we look for sub-classes of terms where this quantification is not mandatory.

Other future works include the study of the behavioral theory of other delimited control operators, like the dynamic ones (e.g., *control* and *prompt* [9] or *shift*$_0$ and *reset*$_0$ [6]), but also of abortive control operators, such as *callcc*, for which no sound and complete bisimilarity has been defined so far.

Acknowledgments We thank Małgorzata Biernacka and the anonymous referees for many helpful comments on the presentation of this work.

References

1. K. Asai and Y. Kameyama. Polymorphic delimited continuations. In Z. Shao, editor, *APLAS'07*, number 4807 in LNCS, pages 239–254, Singapore, Dec. 2007. Springer-Verlag.
2. V. Balat, R. D. Cosmo, and M. P. Fiore. Extensional normalisation and type-directed partial evaluation for typed lambda calculus with sums. In X. Leroy, editor, *POPL'04*, SIGPLAN Notices, Vol. 39, No. 1, pages 64–76, Venice, Italy, Jan. 2004. ACM Press.
3. M. Biernacka, D. Biernacki, and O. Danvy. An operational foundation for delimited continuations in the CPS hierarchy. *Logical Methods in Computer Science*, 1(2:5):1–39, Nov. 2005.
4. D. Biernacki and S. Lenglet. Applicative bisimulations for delimited-control operators. In L. Birkedal, editor, *FOSSACS'12*, number 7213 in LNCS, pages 119–134, Tallinn, Estonia, Mar. 2012. Springer-Verlag.
5. D. Biernacki and S. Lenglet. Normal form bisimulations for delimited-control operators. In T. Schrijvers and P. Thiemann, editors, *FLOPS'12*, number 7294 in LNCS, pages 47–61, Kobe, Japan, May 2012. Springer-Verlag.

6. O. Danvy and A. Filinski. A functional abstraction of typed contexts. DIKU Rapport 89/12, DIKU, Computer Science Department, University of Copenhagen, Copenhagen, Denmark, July 1989.
7. O. Danvy and A. Filinski. Abstracting control. In Wand [23], pages 151–160.
8. R. K. Dybvig, S. Peyton-Jones, and A. Sabry. A monadic framework for delimited continuations. *Journal of Functional Programming*, 17(6):687–730, 2007.
9. M. Felleisen. The theory and practice of first-class prompts. In J. Ferrante and P. Mager, editors, *POPL'88*, pages 180–190, San Diego, California, Jan. 1988. ACM Press.
10. A. Filinski. Representing monads. In H.-J. Boehm, editor, *POPL'94*, pages 446–457, Portland, Oregon, Jan. 1994. ACM Press.
11. R. Hieb, R. K. Dybvig, and C. W. Anderson, III. Subcontinuations. *Lisp and Symbolic Computation*, 5(4):295–326, Dec. 1993.
12. Y. Kameyama. Axioms for control operators in the CPS hierarchy. *Higher-Order and Symbolic Computation*, 20(4):339–369, 2007.
13. Y. Kameyama and M. Hasegawa. A sound and complete axiomatization of delimited continuations. In O. Shivers, editor, *ICFP'03*, SIGPLAN Notices, Vol. 38, No. 9, pages 177–188, Uppsala, Sweden, Aug. 2003. ACM Press.
14. V. Koutavas, P. B. Levy, and E. Sumii. From applicative to environmental bisimulation. *Electronic Notes in Theoretical Computer Science*, 276:215–235, 2011.
15. R. Milner. Fully abstract models of typed λ-calculi. *Theoretical Computer Science*, 4(1):1–22, 1977.
16. J. H. Morris. *Lambda Calculus Models of Programming Languages*. PhD thesis, Massachusets Institute of Technology, 1968.
17. A. Piérard and E. Sumii. A higher-order distributed calculus with name creation. In *LICS'12*, pages 531–540, Dubrovnik, Croatia, June 2012. IEEE Computer Society Press.
18. D. Sangiorgi, N. Kobayashi, and E. Sumii. Environmental bisimulations for higher-order languages. *ACM Transactions on Programming Languages and Systems*, 33(1):1–69, Jan. 2011.
19. N. Sato and E. Sumii. The higher-order, call-by-value applied Pi-calculus. In Z. Hu, editor, *APLAS'09*, volume 5904 of *LNCS*, pages 311–326, Seoul, Korea, Dec. 2009. Springer-Verlag.
20. D. Sitaram and M. Felleisen. Reasoning with continuations II: Full abstraction for models of control. In Wand [23], pages 161–175.
21. E. Sumii. A bisimulation-like proof method for contextual properties in untyped lambda-calculus with references and deallocation. *Theoretical Computer Science*, 411(51-52):4358–4378, 2010.
22. E. Sumii and B. C. Pierce. A bisimulation for dynamic sealing. *Theoretical Computer Science*, 375(1-3):169–192, 2007.
23. M. Wand, editor. *Proceedings of the 1990 ACM Conference on Lisp and Functional Programming*, Nice, France, June 1990. ACM Press.

A Soundness and Completeness for the Relaxed Semantics

In bisimulation up-to environment, one can use bigger environments that the ones needed by Definition 6. As a result, instead of making the environment grow at each bisimulation step, we can directly use the largest possible environment.

Definition 12. *An environmental relation $\mathcal{X}$ is an environmental bisimulation up to environment if*

1. $t_0 \; \mathcal{X}_{\mathcal{E}} \; t_1$ *implies:*
   (a) *if* $t_0 \rightarrow_{\mathrm{v}} t_0'$*, then* $t_1 \rightarrow^*_{\mathrm{v}} t_1'$ *and* $t_0' \; \mathcal{X}_{\mathcal{E}'} \; t_1'$ *for some* $\mathcal{E}'$ *such that* $\mathcal{E} \subseteq \mathcal{E}'$*;*
   (b) *if* t_0 *is a value* v_0*, then* $t_1 \rightarrow^*_{\mathrm{v}} v_1$ *and* $\mathcal{E}' \in \mathcal{X}$ *for some* $\mathcal{E}'$ *such that* $\mathcal{E} \cup \{(v_0, v_1)\} \subseteq \mathcal{E}'$*;*
   (c) *if* t_0 *is a stuck term, then* $t_1 \rightarrow^*_{\mathrm{v}} t_1'$ *where* t_1' *is a stuck term and* $\mathcal{E}' \in \mathcal{X}$ *for some* $\mathcal{E}'$ *such that* $\mathcal{E} \cup \{(t_0, t_1')\} \subseteq \mathcal{E}'$*;*
   (d) *the converse of the above conditions on* t_1*;*
2. $\mathcal{E} \in \mathcal{X}$ *implies:*
   (a) *for all* $(\lambda x.t_0, \lambda x.t_1) \in \mathcal{E}$*, for all* $(v_0, v_1) \in \widehat{\mathcal{E}}$*, we have* $t_0\{v_0/x\} \; \mathcal{X}_{\mathcal{E}'} \; t_1\{v_1/x\}$ *for some* $\mathcal{E} \subseteq \mathcal{E}'$*;*
   (b) *for all* $(E_0[\mathcal{S}k.t_0], E_1[\mathcal{S}k.t_1]) \in \mathcal{E}$*, for all* $(E_0', E_1') \in \widetilde{\mathcal{E}}$*, we have*

$$\langle t_0\{\lambda x.\langle E_0'[E_0[x]]\rangle/k\}\rangle \; \mathcal{X}_{\mathcal{E}'} \; \langle t_1\{\lambda x.\langle E_1'[E_1[x]]\rangle/k\}\rangle$$

   for a fresh x *and some* $\mathcal{E} \subseteq \mathcal{E}'$*.*

Lemma 22. *If* $\mathcal{X}$ *is an environmental bisimulation up to environment, then* $\mathcal{X} \subseteq \approx$*.*

Next, we define bisimulation up-to bisimilarity, where we can compose with $\simeq$ to simplify the definition of candidate relations by factoring out useless bisimilar terms.

Definition 13. *An environmental relation $\mathcal{X}$ is an environmental bisimulation up to bisimilarity if*

1. $t_0 \; \mathcal{X}_{\mathcal{E}} \; t_1$ *implies:*
   (a) *if* $t_0 \rightarrow_{\mathrm{v}} t_0'$*, then* $t_1 \rightarrow^*_{\mathrm{v}} t_1'$ *and* $t_0' \; \mathcal{X}_{\mathcal{E}} \simeq t_1'$*;*
   (b) *if* t_0 *is a value* v_0*, then* $t_1 \rightarrow^*_{\mathrm{v}} v_1$ *and* $\mathcal{E} \cup \{(v_0, v_1')\} \in \mathcal{X}$ *for some* $v_1' \simeq v_1$*;*
   (c) *if* t_0 *is a stuck term, then* $t_1 \rightarrow^*_{\mathrm{v}} t_1'$ *where* t_1' *is a stuck term and* $\mathcal{E} \cup \{(t_0, t_1'')\} \in \mathcal{X}$ *for some stuck term* t_1'' *such that* $t_1'' \simeq t_1'$*;*
   (d) *the converse of the above conditions on* t_1*;*
2. $\mathcal{E} \in \mathcal{X}$ *implies:*
   (a) *for all* $(\lambda x.t_0, \lambda x.t_1) \in \mathcal{E}$*, for all* $(v_0, v_1) \in \widehat{\mathcal{E}}$*, we have* $t_0\{v_0/x\} \simeq \mathcal{X}_{\mathcal{E}} \simeq t_1\{v_1/x\}$*;*

(b) for all $(E_0[\mathcal{S}k.t_0], E_1[\mathcal{S}k.t_1]) \in \mathcal{E}$, *for all* $(E'_0, E'_1) \in \widetilde{\mathcal{E}}$, *we have*

$$\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle \simeq\mathcal{X}_{\mathcal{E}}\simeq \langle t_1\{\lambda x.\langle E'_1[E_1[x]]\rangle/k\}\rangle$$

for a fresh x.

Lemma 23. *If* $\mathcal{X}$ *is an environmental bisimulation up to bisimilarity, then* $\mathcal{X} \subseteq \approx$.

As usual with up-to bisimilarity with small-step relations, we cannot compose on the left-hand side of $\mathcal{X}$ in clause (1) of Definition 13.

Lemma 24. *Let* $\mathcal{R}$ *be a relation on closed terms. If* $t_0\ \widehat{\mathcal{R}}\ t_1$ *(where* t_0 *and* t_1 *are potentially open terms) and* $v_0\ \widehat{\mathcal{R}}^{\mathrm{nf}}\ v_1$, *then* $t_0\{v_0/x\}\ \widehat{\mathcal{R}}\ t_1\{v_1/x\}$.

Proof. We proceed by induction on $t_0\ \widehat{\mathcal{R}}\ t_1$. Suppose $t_0 = t_1 = x$. We have $v_0\ \widehat{\mathcal{R}}^{\mathrm{nf}}\ v_1$ as wished. The result is also easy if $t_0 = t_1 = y \neq x$. Suppose $t_0\ \mathcal{R}\ t_1$. Because $\mathcal{R}$ is defined on closed terms only, we have $t_0\{v_0/x\} = t_0\ \mathcal{R}\ t_1 = t_1\{v_1/x\}$. The remaining induction cases are straightforward.

Lemma 25. *Let* $\mathcal{E}$ *be an environment (i.e., a relation on closed values and closed stuck terms only). Suppose* $t_0\ \widehat{\mathcal{E}}\ t_1$. *If* t_0 *is a value, then so is* t_1, *and if* t_0 *is a stuck term, then so is* t_1.

Proof. The first item is straightforward by case analysis on $t_0\ \widehat{\mathcal{E}}\ t_1$ (and using the fact that $\mathcal{E}$ relates values only with values), and the second item is straightforward by induction on $t_0\ \widehat{\mathcal{E}}\ t_1$ (and using the fact that $\mathcal{E}$ relates stuck terms only with stuck terms).

Lemma 26. *For all* $\mathcal{E}$ *and normal forms* t_0, t_1, *if* $t_0 \approx_{\mathcal{E}} t_1$, *then* $C[t_0] \approx_{\widehat{\mathcal{E}}^{\mathrm{nf}}} C[t_1]$.

Lemma 27. *For all* $\mathcal{E}$, *if* $t_0 \approx_{\mathcal{E}} t_1$, *then* $F[t_0] \approx_{\widehat{\mathcal{E}}^{\mathrm{nf}}} F[t_1]$.

We prove Lemmas 26 and 27 simultaneously. Let $\mathcal{Y}$ be an environmental bisimulation. We define

$$\begin{aligned}
\mathcal{X} &= \mathcal{X}_1 \cup \mathcal{X}_2 \cup \{\widehat{\mathcal{E}}^{\mathrm{nf}} \mid \mathcal{E} \in \mathcal{Y}\}\\
\mathcal{X}_1 &= \{(\widehat{\mathcal{E}}^{\mathrm{nf}}, F_0[t_0], F_1[t_1]) \mid t_0\ \mathcal{Y}_{\mathcal{E}}\ t_1, F_0\ \widetilde{\mathcal{E}}\ F_1\}\\
\mathcal{X}_2 &= \{(\widehat{\mathcal{E}}^{\mathrm{nf}}, t_0, t_1) \mid \mathcal{E} \in \mathcal{Y}, t_0\ \widehat{\mathcal{E}}\ t_1\}
\end{aligned}$$

In $\mathcal{X}_2$, we build the closed terms (t_0, t_1) out of pairs of values or pair of stuck terms. We first prove a preliminary lemma about $\mathcal{X}$.

Lemma 28. *Let* $\mathcal{E} \in \mathcal{Y}$.

– *If* $\lambda x.t_0\ \widehat{\mathcal{E}}\ \lambda x.t_1$ *and* $v_0\ \widehat{\mathcal{E}}^{\mathrm{nf}}\ v_1$ *then* $t_0\{v_0/x\}\ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}}\ t_1\{v_1/x\}$.

– *If* $E_0[\mathcal{S}k.t_0]\ \widehat{\mathcal{E}}\ E_1[\mathcal{S}k.t_1]$ *and* $E'_0\ \widetilde{\mathcal{E}}\ E'_1$, *then* $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle\ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}}\ \langle t_1\{\lambda x.\langle E'_1[E_1[x]]\rangle/k\}\rangle$.

Proof. For the first item, we proceed by case analysis on $\lambda x.t_0\ \widehat{\mathcal{E}}\ \lambda x.t_1$. If $\lambda x.t_0 \mathcal{E} \lambda x.t_1$, then since $\mathcal{Y}$ is an environmental bisimulation, we have $t_0\{v_0/x\}\ \mathcal{Y}_{\mathcal{E}}\ t_1\{v_1/x\}$, which implies $t_0\{v_0/x\}\ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}}\ t_1\{v_1/x\}$ (more precisely, the terms are in $\mathcal{X}_1$).

If $t_0\ \widehat{\mathcal{E}}\ t_1$ with $\mathsf{fv}(t_0) \cup \mathsf{fv}(t_1) \subseteq \{x\}$, then we have $t_0\{v_0/x\}\ \widehat{\mathcal{E}}\ t_1\{v_1/x\}$ by Lemma 24. In fact, we have $t_0\{v_0/x\}\ \widehat{\mathcal{E}}\ t_1\{v_1/x\}$, so we have $t_0\{v_0/x\}\ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}}\ t_1\{v_1/x\}$ (more precisely, the terms are in $\mathcal{X}_2$).

For the second item, we proceed by induction on $E_0[\mathcal{S}k.t_0]\ \widehat{\mathcal{E}}\ E_1[\mathcal{S}k.t_1]$. If $E_0[\mathcal{S}k.t_0] \mathcal{E} E_1[\mathcal{S}k.t_1]$, then because $\mathcal{Y}$ is an environmental bisimulation, we have $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle\ \mathcal{Y}_{\mathcal{E}}\ \langle t_1\{\lambda x.\langle E'_1[E_1[x]]\rangle/k\}\rangle$, which is equivalent to $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle\ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}}\ \langle t_1\{\lambda x.\langle E'_1[E_1[x]]\rangle/k\}\rangle$ (the terms are in $\mathcal{X}_1$).

Suppose $E_0 = E_1 = \square$ and $t_0\ \widehat{\mathcal{E}}\ t_1$ with $\mathsf{fv}(t_0) \cup \mathsf{fv}(t_1) \subseteq \{k\}$. From $E'_0\ \widetilde{\mathcal{E}}\ E'_1$, we deduce $\lambda x.\langle E'_0[x]\rangle\ \widehat{\mathcal{E}}\ \lambda x.\langle E'_1[x]\rangle$. We have $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle\ \widehat{\mathcal{E}}\ \langle t_1\{\lambda x.\langle E'_1[E_1[x]]\rangle/k\}\rangle$, by Lemma 24, hence the result holds (the terms are in $\mathcal{X}_2$).

Suppose $E_0 = v_0\ E''_0$ and $E_1 = v_1\ E''_1$ with $v_0\ \widehat{\mathcal{E}}\ v_1$ and $E''_0[\mathcal{S}k.t_0]\ \widehat{\mathcal{E}}\ E''_1[\mathcal{S}k.t_1]$. From $v_0\ \widehat{\mathcal{E}}\ v_1$ and $E'_0\ \widetilde{\mathcal{E}}\ E'_1$, we deduce $E'_1[v_0\ \square]\ \widetilde{\mathcal{E}}\ v'_1[v_1\ \square]$. Then $\langle t_0\{\lambda x.\langle E'_0[v_0\ E''_0[x]]\rangle/k\}\rangle\ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}}\ \langle t_1\{\lambda x.\langle E'_1[v_1\ E''_1[x]]\rangle/k\}\rangle$ by the induction hypothesis, i.e., $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle\ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}}\ \langle t_1\{\lambda x.\langle E'_1[E_1[x]]\rangle/k\}\rangle$, as wished. The case $E_0 = E''_0\ t'_0$ and $E_1 = E''_1\ t'_1$ is similar.

We now prove Lemmas 26 and 27 by showing that $\mathcal{X}$ is a bisimulation up to environment.

Proof. We first prove the bisimulation for the elements in $\mathcal{X}_2$ (for these, we do not need the "up to environment"). Let $t_0\ \widehat{\mathcal{E}}\ t_1$, with $\mathcal{E} \in \mathcal{Y}$. Clause 1b (resp. 1c) is easy: if t_0 is a value (resp. a stuck term), then so is t_1 (cf. Lemma 25), and we have $\widehat{\mathcal{E}}^{\mathrm{nf}} \cup \{(t_0, t_1)\} = \widehat{\mathcal{E}}^{\mathrm{nf}} \in \mathcal{X}$. For clause 1a, we proceed by induction on $t_0\ \widehat{\mathcal{E}}\ t_1$.

Suppose $t_0 = t^1_0\ t^2_0$ and $t_1 = t^1_1\ t^2_1$ with $t^1_0\ \widehat{\mathcal{E}}\ t^1_1$ and $t^2_0\ \widehat{\mathcal{E}}\ t^2_1$. We have three cases to consider.

- Assume $t^1_0 \rightarrow_{\mathrm{v}} t^{1\prime}_0$, so that $t_0 \rightarrow_{\mathrm{v}} t^{1\prime}_0\ t^2_0$. By the induction hypothesis, there exists $t^{1\prime}_1$ such that $t^1_1 \rightarrow^*_{\mathrm{v}} t^{1\prime}_1$ and $t^{1\prime}_0\ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}}\ t^{1\prime}_1$. From $t^2_0\ \widehat{\mathcal{E}}\ t^2_1$ and $t^{1\prime}_0\ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}}\ t^{1\prime}_1$, we can deduce $t^{1\prime}_0\ t^2_0\ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}}\ t^{1\prime}_1\ t^2_1$ by definition of $\mathcal{X}$. We also have $t_1 \rightarrow^*_{\mathrm{v}} t^{1\prime}_1\ t^2_1$, hence the result holds.
- Assume $t^1_0 = v_0$ and $t^2_0 \rightarrow_{\mathrm{v}} t^{2\prime}_0$, so that $t_0 \rightarrow_{\mathrm{v}} v_0\ t^{2\prime}_0$. Then t^1_1 is also a value v_1 according to Lemma 25. By the induction hypothesis, there exists $t^{2\prime}_1$ such that $t^2_1 \rightarrow^*_{\mathrm{v}} t^{2\prime}_1$ and $t^{2\prime}_0\ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}}\ t^{2\prime}_1$. From $v_0\ \widehat{\mathcal{E}}\ v^2_1$ and $t^{2\prime}_0\ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}}\ t^{2\prime}_1$, we can deduce $v_0\ t^{2\prime}_0\ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}}\ v_1\ t^{2\prime}_1$ by definition of $\mathcal{X}$. We also have $t_1 \rightarrow^*_{\mathrm{v}} v_1\ t^{2\prime}_1$, hence the result holds.

– Assume $t_0^1 = \lambda x.t_0'$ and $t_0^2 = v_0$, so that $t_0 \rightarrow_{\mathrm{v}} t_0'\{v_0/x\}$. By Lemma 25, t_1^1 is a value $\lambda x.t_1'$ and t_1^2 is a value v_1. We have $t_1 \rightarrow_{\mathrm{v}} t_0'\{v_0/x\}$, and by Lemma 28, we have $t_0'\{v_0/x\} \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}} \ t_1'\{v_1/x\}$, hence the result holds.

Suppose $t_0 = \langle t_0' \rangle$, $t_1 = \langle t_1' \rangle$ with $t_0' \ \widehat{\mathcal{E}} \ t_1'$. We have two possibilities.

– Assume $t_0' \rightarrow_{\mathrm{v}} t_0''$, so that $t_0 \rightarrow_{\mathrm{v}} \langle t_0'' \rangle$. By the induction hypothesis, there exists t_1'' such that $t_1' \rightarrow_{\mathrm{v}}^* t_1''$ and $t_0'' \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}} \ t_1''$. By definition of $\mathcal{X}$, we have $\langle t_0'' \rangle \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}} \ \langle t_1'' \rangle$, and furthermore $t_0 \rightarrow_{\mathrm{v}}^* \langle t_1'' \rangle$, we therefore have the required result.
– Assume $t_0' = E_0[\mathcal{S}k.t_0'']$, so that $t_0 \rightarrow_{\mathrm{v}} \langle t_0''\{\lambda x.\langle E_0[x] \rangle/k\} \rangle$. By Lemma 25, t_1' is a stuck term $E_1[\mathcal{S}k.t_1'']$, therefore $t_1 \rightarrow_{\mathrm{v}} \langle t_1''\{\lambda x.\langle E_1[x] \rangle/k\} \rangle$. We have $\langle t_0''\{\lambda x.\langle E_0[x] \rangle/k\} \rangle \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}} \ \langle t_1''\{\lambda x.\langle E_1[x] \rangle/k\} \rangle$ by Lemma 28, hence the result holds.

We now prove the bisimulation property (up to environment) for elements in $\mathcal{X}_1$. Let $F_0[t_0] \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}} \ F_1[t_1]$, so that $t_0 \ \mathcal{Y}_{\mathcal{E}} \ t_1$ and $F_0 \ \widetilde{\mathcal{E}} \ F_1$. If t_0 is a value v_0, then because $\mathcal{Y}$ is a bisimulation, there exists v_1 such that $t_1 \rightarrow_{\mathrm{v}}^* v_1$ and $\mathcal{E}' = \mathcal{E} \cup \{(t_0, t_1)\} \in \mathcal{Y}$. We then have $F_1[t_1] \rightarrow_{\mathrm{v}}^* F_1[v_1]$, and the terms $F_0[v_0]$, $F_1[v_1]$ are similar to the one of $\mathcal{X}_1$. We can prove the bisimulation property with $F_0[v_0]$, $F_1[v_1]$ the same way we did with the terms in $\mathcal{X}_1$, except that we reason up to environment, because $\mathcal{E} \subseteq \mathcal{E}'$. The reasoning is similar if t_0 is a stuck term. Suppose t_0 is not a value nor a stuck term. There exists t_0' such that $t_0 \rightarrow_{\mathrm{v}} t_0'$, and so $F_0[t_0] \rightarrow_{\mathrm{v}} F_0[t_0']$. Because $\mathcal{Y}$ is a bisimulation, there exists t_1' such that $t_1 \rightarrow_{\mathrm{v}}^* t_1'$ and $t_0' \ \mathcal{Y}_{\mathcal{E}} \ t_1'$. We therefore have $F_1[t_1] \rightarrow_{\mathrm{v}}^* F_1[t_1']$ with $F_0[t_0'] \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}} \ F_1[t_1']$, as wished.

We now prove the clause 2 of the bisimulation. The only environments in $\mathcal{X}$ are of the form $\widehat{\mathcal{E}}^{\mathrm{nf}}$. Let $\lambda x.t_0 \ \widehat{\mathcal{E}}^{\mathrm{nf}} \ \lambda x.t_1$ and $v_0 \ \widehat{\mathcal{E}}^{\mathrm{nf}} \ v_1$. By Lemma 28, we have $t_0\{v_0/x\} \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}} \ t_1\{v_1/x\}$, hence the result holds. Similarly, if $E_0[\mathcal{S}k.t_0] \ \widehat{\mathcal{E}}^{\mathrm{nf}} \ E_1[\mathcal{S}k.t_1]$ and $E_0' \ \widetilde{\mathcal{E}} \ E_1'$, then (by Lemma 28) we have $\langle t_0\{\lambda x.\langle E_0'[E_0[x]] \rangle/k\} \rangle \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{nf}}} \ \langle t_1\{\lambda x.\langle E_1'[E_1[x]] \rangle/k\} \rangle$.

Lemma 29. *If $\lambda x.t_0 \simeq \lambda x.t_1$, then $t_0\{v/x\} \simeq t_1\{v/x\}$.*

Proof. By clause 1b, we have $\{(\lambda x.t_0, \lambda x.t_1)\} \in \simeq$. Let $\mathcal{E} = \{(\lambda x.t_0, \lambda x.t_1)\}$. By clause 2a, for all v, we have $t_0\{v/x\} \approx_{\mathcal{E}} t_1\{v/x\}$, therefore $t_0\{v/x\} \simeq t_1\{v/x\}$ holds by weakening (Lemma 5).

Lemma 30. *If $E_0[\mathcal{S}k.t_0] \simeq E_1[\mathcal{S}k.t_1]$, then we have $\langle t_0\{\lambda x.\langle E[E_0[x]] \rangle/k\} \rangle \simeq \langle t_1\{\lambda x.\langle E[E_1[x]] \rangle/k\} \rangle$.*

Proof. By clause 1c, we know that $\{(E_0[\mathcal{S}k.t_0], E_1[\mathcal{S}k.t_1])\} \in \simeq$. Let $\mathcal{E} = \{(E_0[\mathcal{S}k.t_0], E_1[\mathcal{S}k.t_1])\}$. By clause 2b, we know that $\langle t_0\{\lambda x.\langle E[E_0[x]] \rangle/k\} \rangle \approx_{\mathcal{E}} \langle t_1\{\lambda x.\langle E[E_1[x]] \rangle/k\} \rangle$, hence $\langle t_0\{\lambda x.\langle E[E_0[x]] \rangle/k\} \rangle \simeq \langle t_1\{\lambda x.\langle E[E_1[x]] \rangle/k\} \rangle$ is true by weakening (Lemma 5).

Lemma 31. *If $\lambda x.t_0 \ \widehat{\simeq} \ \lambda x.t \simeq \lambda x.t_1$ and $v_0 \ \widehat{\simeq} \ v \simeq v_1$ then $t_0\{v_0/x\} \ \widehat{\simeq}\simeq \ t_1\{v_1/x\}$.*

Proof. We proceed by case analysis on $\lambda x.t_0 \mathrel{\widehat{\simeq}} \lambda x.t$.

Suppose $\lambda x.t_0 \simeq \lambda x.t$. We have $t_0\{v_0/x\} \mathrel{\widehat{\simeq}} t_0\{v/x\}$ by Lemma 24, $t_0\{v/x\} \simeq t\{v/x\}$ by Lemma 29, $t\{v/x\} \simeq t\{v_1/x\}$ by Lemma 26, and $t\{v_1/x\} \simeq t_1\{v_1/x\}$ by Lemma 29. Finally, $t_0\{v_0/x\} \mathrel{\widehat{\simeq}\simeq} t_1\{v_1/x\}$ holds using transitivity of $\simeq$.

Suppose $t_0 \mathrel{\widehat{\simeq}} t$ with $\mathsf{fv}(t_0) \cup \mathsf{fv}(t) \subseteq \{x\}$. We have $t_0\{v_0/x\} \mathrel{\widehat{\simeq}} t\{v/x\}$ by Lemma 24, $t\{v/x\} \simeq t\{v_1/x\}$ by Lemma 26, and $t\{v_1/x\} \simeq t_1\{v_1/x\}$ by Lemma 29. Finally, $t_0\{v_0/x\} \mathrel{\widehat{\simeq}\simeq} t_1\{v_1/x\}$ holds using transitivity of $\simeq$.

Lemma 32. *If $E_0[\mathcal{S}k.t_0] \mathrel{\widehat{\simeq}} E[\mathcal{S}k.t]$, $E[\mathcal{S}k.t] \simeq E_1[\mathcal{S}k.t_1]$ and $E'_0 \mathrel{\widehat{\simeq}} E'_1$, then $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle \mathrel{\widehat{\simeq}\simeq} \langle t_1\{\lambda x.\langle E'_1[E_1[x]]\rangle/k\}\rangle$.*

Proof. We start with proving that $E_0[\mathcal{S}k.t_0] \mathrel{\widehat{\simeq}} E[\mathcal{S}k.t]$ and $E'_0 \mathrel{\widehat{\simeq}} E'_1$ implies $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle \mathrel{\widehat{\simeq}\simeq} \langle t\{\lambda x.\langle E'_1[E[x]]\rangle/k\}\rangle$. We proceed by induction on $E_0[\mathcal{S}k.t_0] \mathrel{\widehat{\simeq}} E[\mathcal{S}k.t]$.

Suppose $E_0[\mathcal{S}k.t_0] \simeq E[\mathcal{S}k.t]$. From $E'_0 \mathrel{\widehat{\simeq}} E'_1$, we get $\lambda x.\langle E'_0[E_0[x]]\rangle \mathrel{\widehat{\simeq}} \lambda x.\langle E'_1[E_0[x]]\rangle$. Then $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle \mathrel{\widehat{\simeq}} \langle t_0\{\lambda x.\langle E'_1[E_0[x]]\rangle/k\}\rangle$ holds by Lemma 24, and then $\langle t_0\{\lambda x.\langle E'_1[E_0[x]]\rangle/k\}\rangle \simeq \langle t\{\lambda x.\langle E'_1[E[x]]\rangle/k\}\rangle$ holds by Lemma 30, hence the result holds.

Suppose $E_0 = E = \square$ and $t_0 \mathrel{\widehat{\simeq}} t$ with $\mathsf{fv}(t_0) \cup \mathsf{fv}(t) \subseteq \{k\}$. From $E'_0 \mathrel{\widehat{\simeq}} E'_1$, we have $\lambda x.\langle E'_0[x]\rangle \mathrel{\widehat{\simeq}} \lambda x.\langle E'_1[x]\rangle$. Then $\langle t_0\{\lambda x.\langle E'_0[x]\rangle/k\}\rangle \mathrel{\widehat{\simeq}} \langle t\{\lambda x.\langle E'_1[x]\rangle/k\}\rangle$ by Lemma 24, hence the result holds.

Suppose $E_0[\mathcal{S}k.t_0] = v_0\ E''_0[\mathcal{S}k.t_0]$, $E[\mathcal{S}k.t] = v\ E''[\mathcal{S}k.t]$ with $v_0 \mathrel{\widehat{\simeq}} v$ and $E''_0[\mathcal{S}k.t_0] \mathrel{\widehat{\simeq}} E''[\mathcal{S}k.t]$. From $E'_0 \mathrel{\widehat{\simeq}} E'_1$ and $v_0 \mathrel{\widehat{\simeq}} v$, it is the case that $E'_0[v_0\ \square] \mathrel{\widehat{\simeq}} E'_1[v\ \square]$. By the induction hypothesis, we obtain

$$\langle t_0\{\lambda x.\langle E'_0[v_0\ E''_0[x]]\rangle/k\}\rangle \mathrel{\widehat{\simeq}\simeq} \langle t\{\lambda x.\langle E'_1[v\ E[x]]\rangle/k\}\rangle,$$

which means that $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle \mathrel{\widehat{\simeq}\simeq} \langle t\{\lambda x.\langle E'_1[E[x]]\rangle/k\}\rangle$, as wished. The other case $E_0[\mathcal{S}k.t_0] = E''_0[\mathcal{S}k.t_0]\ t'_0$, $E[\mathcal{S}k.t] = E''[\mathcal{S}k.t]\ t'$ with $t'_0 \mathrel{\widehat{\simeq}} t'$ and $E''_0[\mathcal{S}k.t_0] \mathrel{\widehat{\simeq}} E''[\mathcal{S}k.t]$ is treated similarly.

We are now in a position to prove the lemma. We have just proved that $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle \mathrel{\widehat{\simeq}\simeq} \langle t\{\lambda x.\langle E'_1[E[x]]\rangle/k\}\rangle$. We also have that

$$\langle t\{\lambda x.\langle E'_1[E[x]]\rangle/k\}\rangle \simeq \langle t_1\{\lambda x.\langle E'_1[E_1[x]]\rangle/k\}\rangle$$

by Lemma 30, therefore the required result holds by transitivity of $\simeq$.

Lemma 33. *$t_0 \simeq t_1$ implies $C[t_0] \approx_{\widehat{\simeq}^{\mathrm{nf}}} C[t_1]$.*

Proof. We prove that

$$\mathcal{X} = \{(\widehat{\simeq}^{\mathrm{nf}}, t_0, t_1) \mid t_0 \mathrel{\widehat{\simeq}} t_1\} \cup \{\widehat{\simeq}^{\mathrm{nf}}\}$$

is a bisimulation up-to bisimilarity. Let $t_0 \mathrel{\mathcal{X}_{\widehat{\simeq}^{\mathrm{nf}}}} t_1$. We prove clauses 1a, 1b, and 1c of Definition 13 by induction on $t_0 \mathrel{\widehat{\simeq}} t_1$. Note that by definition of $\mathcal{X}$, we have $t \mathrel{\mathcal{X}_{\widehat{\simeq}^{\mathrm{nf}}}} t'$ iff $t \mathrel{\widehat{\simeq}} t'$.

Suppose $t_0 \simeq t_1$. This case holds because $\simeq$ is an environmental bisimulation.

Suppose $t_0 = \lambda x.t'_0$, $t_1 = \lambda x.t'_1$ with $t'_0 \mathrel{\widehat{\simeq}} t'_1$ and $\mathsf{fv}(t'_0) \cup \mathsf{fv}(t'_1) \subseteq \{x\}$. We have to prove that $(\widehat{\simeq}^{\mathrm{nf}} \cup \{(t_0, t_1)\}) \in \mathcal{X}$, i.e., $\widehat{\simeq}^{\mathrm{nf}} \in \mathcal{X}$, which is true.

Suppose $t_0 = t_0^1\, t_0^2$, $t_1 = t_1^1\, t_0^2$ with $t_0^1 \mathrel{\widehat{\simeq}} t_1^1$ and $t_0^2 \mathrel{\widehat{\simeq}} t_1^2$. We distinguish several cases.

- If $t_0^1 \rightarrow_{\mathrm{v}} {t_0^1}'$, then $t_0 \rightarrow_{\mathrm{v}} {t_0^1}'\, t_0^2$. By induction there exist ${t_1^1}''$, ${t_1^1}'$ such that $t_1^1 \rightarrow_{\mathrm{v}}^* {t_1^1}'$ and ${t_0^1}' \mathrel{\widehat{\simeq}} {t_1^1}'' \simeq {t_1^1}'$. Consequently we have $t_1 \rightarrow_{\mathrm{v}}^* {t_1^1}'\, t_1^2$. By definition, we have ${t_0^1}'\, t_0^2 \mathrel{\widehat{\simeq}} {t_1^1}''\, t_1^2$, and by Lemma 27, we have ${t_1^1}''\, t_1^2 \simeq {t_1^1}'\, t_1^2$, hence ${t_0^1}'\, t_0^2 \mathrel{\widehat{\simeq}\simeq} {t_1^1}'\, t_1^2$ holds, as wished.
- If $t_0^1 = v_0$ and $t_0^2 \rightarrow_{\mathrm{v}} {t_0^2}'$, then $t_0 \rightarrow_{\mathrm{v}} v_0\, {t_0^2}'$. By induction there exist ${t_1^2}''$, ${t_1^2}'$ such that $t_1^2 \rightarrow_{\mathrm{v}}^* {t_1^2}'$ and ${t_0^2}' \mathrel{\widehat{\simeq}} {t_1^2}'' \simeq {t_1^2}'$. There also exists v'_1, v_1 such that $t_1^1 \rightarrow_{\mathrm{v}}^* v_1$ and $v_0 \mathrel{\widehat{\simeq}^{\mathrm{nf}}} v'_1 \simeq v_1$. Consequently we have $t_1 \rightarrow_{\mathrm{v}}^* v_1\, {t_1^2}'$. By definition, we have $v_0\, {t_0^2}' \mathrel{\widehat{\simeq}} v'_1\, {t_1^2}''$, and by Lemma 27 and transitivity of $\simeq$, we have $v'_1\, {t_1^2}'' \simeq v_1\, {t_1^2}'$, hence $v_0\, {t_0^2}' \mathrel{\widehat{\simeq}\simeq} v_1\, {t_1^2}'$ holds, as wished.
- If $t_0^1 = \lambda x.t'_0$ and $t_0^2 \rightarrow_{\mathrm{v}} v_0$, then $t_0 \rightarrow_{\mathrm{v}} t'_0\{v_0/x\}$. By induction there exist t''_1, t'_1 such that $t_1^1 \rightarrow_{\mathrm{v}}^* \lambda x.t'_1$ and $\lambda x.t'_0 \mathrel{\widehat{\simeq}^{\mathrm{nf}}} \lambda x.t''_1 \simeq \lambda x.t'_1$. There also exists v'_1, v_1 such that $t_1^2 \rightarrow_{\mathrm{v}}^* v_1$ and $v_0 \mathrel{\widehat{\simeq}^{\mathrm{nf}}} v'_1 \simeq v_1$. Consequently we have $t_1 \rightarrow_{\mathrm{v}}^* t'_1\{v_1/x\}$. By Lemma 31, we have $t'_0\{v_0/x\} \mathrel{\widehat{\simeq}\simeq} t'_1\{v_1/x\}$, as wished.
- If $t_0 = E_0[\mathcal{S}k.t'_0]\, t_0^2$, then by induction there exist E_1 and t'_1 such that $t_1^1 \rightarrow_{\mathrm{v}}^* E_1[\mathcal{S}k.t'_1]$ and $E_0[\mathcal{S}k.t'_0] \mathrel{\widehat{\simeq}\simeq} E_1[\mathcal{S}k.t'_1]$. By definition of $\widehat{\simeq}$ and Lemma 27, we have $E_0[\mathcal{S}k.t'_0]\, t_0^2 \mathrel{\widehat{\simeq}\simeq} E_1[\mathcal{S}k.t'_1]\, t_1^2$, hence the result holds. The reasoning is the same if $t_0 = v_0\, E_0[\mathcal{S}k.t'_0]$.

Suppose $t_0 = \langle t'_0 \rangle$ and $t_1 = \langle t'_1 \rangle$ with $t'_0 \mathrel{\widehat{\simeq}} t'_1$. We have three cases to consider.

- If $t'_0 \rightarrow_{\mathrm{v}} t''_0$, then $t_0 \rightarrow_{\mathrm{v}} \langle t''_0 \rangle$. By induction there exists t''_1 such that $t'_1 \rightarrow_{\mathrm{v}}^* t''_1$ and $t''_0 \mathrel{\widehat{\simeq}\simeq} t''_1$. Consequently we have $t_1 \rightarrow_{\mathrm{v}}^* \langle t''_1 \rangle$, and by definition of $\widehat{\simeq}$ and Lemma 27, we have $\langle t''_0 \rangle \mathrel{\widehat{\simeq}\simeq} \langle t''_1 \rangle$.
- If $t'_0 = E_0[\mathcal{S}k.t''_0]$, then $t_0 \rightarrow_{\mathrm{v}} \langle t''_0\{\lambda x.\langle E_0[x]\rangle/k\}\rangle$. By induction, there exist E_1 and t''_1 such that $t'_1 \rightarrow_{\mathrm{v}}^* E_1[\mathcal{S}k.t''_1]$ and $E_0[\mathcal{S}k.t''_0] \mathrel{\widehat{\simeq}\simeq} E_1[\mathcal{S}k.t''_1]$. By Lemma 32, we have $\langle t''_0\{\lambda x.\langle E_0[x]\rangle/k\}\rangle \mathrel{\widehat{\simeq}\simeq} \langle t''_1\{\lambda x.\langle E_1[x]\rangle/k\}\rangle$, as wished.
- If $t'_0 = v_0$, then $t_0 \rightarrow_{\mathrm{v}} v_0$. By induction, there exists v_1 such that $t'_1 \rightarrow_{\mathrm{v}}^* v_1$ and $v_0 \mathrel{\widehat{\simeq}\simeq} v_1$. We have $t_1 \rightarrow_{\mathrm{v}}^* v_1$, hence the result holds.

Suppose $t_0 = \mathcal{S}k.t'_0$ and $t_1 = \mathcal{S}k.t'_1$ with $t'_0 \mathrel{\widehat{\simeq}} t'_1$ and $\mathsf{fv}(t'_0) \cup \mathsf{fv}(t'_1) \subseteq \{x\}$. We have to prove that $(\widehat{\simeq}^{\mathrm{nf}} \cup \{(t_0, t_1)\}) \in \mathcal{X}$, i.e., $\widehat{\simeq}^{\mathrm{nf}} \in \mathcal{X}$, which is true.

We now prove items 2a and 2b of Definition 13. Suppose $\lambda x.t_0 \mathrel{\widehat{\simeq}} \lambda x.t_1$ and $v_0 \mathrel{\widehat{\simeq}} v_1$. Then by Lemma 31 and reflexivity of $\simeq$, we have $t_0\{v_0/x\} \mathrel{\simeq\widehat{\simeq}\simeq} t_1\{v_0/x\}$, as wished.

Suppose $E_0[\mathcal{S}k.t_0] \mathrel{\widehat{\simeq}} E_1[\mathcal{S}k.t_1]$ and $E'_0 \mathrel{\widehat{\simeq}} E'_1$. Then by Lemma 32 and reflexivity of $\simeq$, $\langle t_0\{\lambda x.\langle E'_0[E_0[x]]\rangle/k\}\rangle \mathrel{\simeq\widehat{\simeq}\simeq} \langle t_1\{\lambda x.\langle E'_1[E_1[x]]\rangle/k\}\rangle$, as wished.

Remark 2. The proof of Lemma 33 uses up to bisimilarity because of Lemma 31, and Lemma 31 cannot be strengthened.

Corollary 4. *For all $\mathcal{E} \in \approx$, $\approx_{\mathcal{E}}$ is a congruence.*

Proof. If $t_0 \approx_{\mathcal{E}} t_1$, then by weakening (Lemma 5), we have $t_0 \simeq t_1$, which in turn implies $C[t_0] \approx_{\simeq^{\mathrm{nf}}} C[t_1]$ (by Lemma 33), and gives us $C[t_0] \approx_{\mathcal{E}} C[t_1]$ using weakening again.

Corollary 5 (Soundness). *The relation $\simeq$ is sound.*

Proof. Because it is a congruence, and the observables actions coincide.

Theorem 2 (Completeness). *The relation $\simeq$ is complete.*

Proof. We prove that $\mathcal{X} = \{(\approx_c{}^{\mathrm{nf}}, t_0, t_1) \mid t_0 \approx_c t_1\} \cup \{\approx_c{}^{\mathrm{nf}}\}$ is a big-step environmental bisimulation.

Let $t_0 \; \mathcal{X}_{\approx_c{}^{\mathrm{nf}}} \; t_1$. If $t_0 \rightarrow_{\mathrm{v}}^* v_0$, then by definition of $\approx_c$, there exists v_1 such that $t_1 \rightarrow_{\mathrm{v}}^* v_1$. By Lemma 6, we have $t_0 \simeq v_0$ and $t_1 \simeq v_1$, which implies $t_0 \approx_c v_0$ and $t_1 \approx_c v_1$ by Corollary 5. Transitivity of $\approx_c$ gives $v_0 \approx_c v_1$, hence we have $\approx_c{}^{\mathrm{nf}} \cup \{(v_0, v_1)\} = \approx_c{}^{\mathrm{nf}} \in \mathcal{X}$, as wished. The reasoning is the same for $t_0 \rightarrow_{\mathrm{v}}^* t_0'$, where t_0' is a stuck term.

Let $\lambda x.t_0 \approx_c \lambda x.t_1$ and $v_0 \; \widehat{\approx_c{}^{\mathrm{nf}}} \; v_1$. By congruence of $\approx_c$, we have $v_0 \approx_c v_1$, and also $(\lambda x.t_0)\, v_0 \approx_c (\lambda x.t_1)\, v_1$. Because $(\lambda x.t_0) \rightarrow_{\mathrm{v}} t_0\{v_0/x\}$, $(\lambda x.t_1)\, v_1 \rightarrow_{\mathrm{v}} t_1\{v_1/x\}$ and $\rightarrow_{\mathrm{v}} \subseteq \simeq \subseteq \approx_c$, we have $t_0\{v_0/x\} \approx_c t_1\{v_1/x\}$, i.e., $t_0\{v_0/x\} \; \mathcal{X}_{\approx_c{}^{\mathrm{nf}}} \; t_1\{v_1/x\}$, as wished.

Let $E_0[\mathcal{S}k.t_0] \approx_c E_1[\mathcal{S}k.t_1]$ and $E_0' \; \widetilde{\approx_c{}^{\mathrm{nf}}} \; E_1'$. By induction on $E_0' \; \widetilde{\approx_c{}^{\mathrm{nf}}} \; E_1'$, we know that $\langle E_0'[E_0[\mathcal{S}k.t_0]]\rangle \approx_c \langle E_1'[E_1[\mathcal{S}k.t_1]]\rangle$ holds. From $\langle E_0'[E_0[\mathcal{S}k.t_0]]\rangle \rightarrow_{\mathrm{v}} \langle t_0\{\lambda x.\langle E_0'[E_0[x]]\rangle/k\}\rangle$, $\langle E_1'[E_1[\mathcal{S}k.t_1]]\rangle \rightarrow_{\mathrm{v}} \langle t_1\{\lambda x.\langle E_1'[E_1[x]]\rangle/k\}\rangle$, and $\rightarrow_{\mathrm{v}} \subseteq \simeq \subseteq \approx_c$, we get $\langle t_0\{\lambda x.\langle E_0'[E_0[x]]\rangle/k\}\rangle \approx_c \langle t_1\{\lambda x.\langle E_1'[E_1[x]]\rangle/k\}\rangle$, and then, as required, $\langle t_0\{\lambda x.\langle E_0'[E_0[x]]\rangle/k\}\rangle \; \mathcal{X}_{\approx_c{}^{\mathrm{nf}}} \; \langle t_1\{\lambda x.\langle E_1'[E_1[x]]\rangle/k\}\rangle$.

B Soundness and Completeness for the Original Semantics

Lemma 34. *If $t_0 \dot{\approx}_{\mathcal{E}} t_1$ and $\mathcal{E}' \subseteq \mathcal{E}$ then $t_0 \dot{\approx}_{\mathcal{E}'} t_1$.*

Proof. As in [18].

Lemma 35. *If $t_0 \rightarrow_{\mathrm{v}} t_0'$, then $t_0 \dot{\simeq} t_0'$.*

Proof. Same as for Lemma 6.

Lemma 36. *For all $\mathcal{E}$, if $v_0 \approx_{\mathcal{E}} v_1$, then $C[v_0] \dot{\approx}_{\widehat{\mathcal{E}}^{\mathrm{v}}} C[v_1]$.*

Lemma 37. *For all $\mathcal{E}$, if $t_0 \approx_{\mathcal{E}} t_1$, then $F[t_0] \dot{\approx}_{\widehat{\mathcal{E}}^{\mathrm{v}}} F[t_1]$.*

We prove Lemmas 36 and 37 simultaneously. Let $\mathcal{Y}$ be an environmental bisimulation. We define

$$\begin{aligned} \mathcal{X} &= \mathcal{X}_1 \cup \mathcal{X}_2 \cup \{\widehat{\mathcal{E}}^{\mathrm{v}} \mid \mathcal{E} \in \mathcal{Y}\} \\ \mathcal{X}_1 &= \{(\widehat{\mathcal{E}}^{\mathrm{v}}, F_0[t_0], F_1[t_1]) \mid t_0 \; \mathcal{Y}_{\mathcal{E}} \; t_1, F_0 \; \widetilde{\mathcal{E}} \; F_1\} \\ \mathcal{X}_2 &= \{(\widehat{\mathcal{E}}^{\mathrm{v}}, t_0, t_1) \mid \mathcal{E} \in \mathcal{Y}, t_0 \; \widehat{\mathcal{E}} \; t_1\} \end{aligned}$$

In $\mathcal{X}_2$, we build the closed terms (t_0, t_1) out of pairs of values. We first prove a preliminary lemma about $\mathcal{X}$. Remark that $\mathcal{X}$ is a congruence.

Lemma 38. *Let $\mathcal{E} \in \mathcal{Y}$. If $\lambda x.t_0 \ \widehat{\mathcal{E}} \ \lambda x.t_1$ and $v_0 \ \widehat{\mathcal{E}}^{\mathrm{v}} \ v_1$ then $t_0\{v_0/x\} \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{v}}} \ t_1\{v_1/x\}$.*

Proof. We proceed by case analysis on $\lambda x.t_0 \ \widehat{\mathcal{E}} \ \lambda x.t_1$. If $\lambda x.t_0 \mathcal{E} \lambda x.t_1$, then because $\mathcal{Y}$ is an environmental bisimulation, we have $t_0\{v_0/x\} \ \mathcal{Y}_{\mathcal{E}} \ t_1\{v_1/x\}$, which implies $t_0\{v_0/x\} \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{v}}} \ t_1\{v_1/x\}$ (more precisely, the terms are in $\mathcal{X}_1$).

If $t_0 \ \widehat{\mathcal{E}} \ t_1$ with $\mathsf{fv}(t_0) \cup \mathsf{fv}(t_1) \subseteq \{x\}$, then we have $t_0\{v_0/x\} \ \widehat{\mathcal{E}} \ t_1\{v_1/x\}$ by Lemma 24. In fact, we have $t_0\{v_0/x\} \ \widehat{\mathcal{E}} \ t_1\{v_1/x\}$, so we have $t_0\{v_0/x\} \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{v}}} \ t_1\{v_1/x\}$ (more precisely, the terms are in $\mathcal{X}_2$).

We now prove Lemmas 36 and 37 by showing that $\mathcal{X}$ is a bisimulation up to environment.

Proof. We first prove the bisimulation for the elements in $\mathcal{X}_2$ (for these, we do not need the "up to environment"). Let $t_0 \ \widehat{\mathcal{E}} \ t_1$, with $\mathcal{E} \in \mathcal{Y}$. If t_0 or t_1 is not a program, then for all $E_0 \ \widetilde{\mathcal{E}} \ E_1$, we have $\langle E_0[t_0]\rangle \ \widehat{\mathcal{E}} \ \langle E_1[t_1]\rangle$, i.e., $\langle E_0[t_0]\rangle \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{v}}} \ \langle E_1[t_1]\rangle$, hence clause 1 holds. Suppose t_0, t_1 are programs p_0, p_1. If $p_0 \rightarrow_{\mathrm{v}} v_0$, then $p_0 = \langle v_0\rangle$, and therefore $p_1 = \langle v_1\rangle$ with $v_0 \ \widehat{\mathcal{X}} \ v_1$. We have $p_1 \rightarrow_{\mathrm{v}} v_1$, and also $\{(v_0, v_1)\} \cup \widehat{\mathcal{E}}^{\mathrm{v}} = \widehat{\mathcal{E}}^{\mathrm{v}} \in \mathcal{X}$, as wished by clause 2b.

Otherwise $p_0 \rightarrow_{\mathrm{v}} p_0'$. Then $p_0 = F_0[r_0]$. Because $\mathcal{E}$ relates only values, we can prove there exist F_1, r_1 such that $p_1 = F_1[r_1]$, $F_0 \ \widetilde{\mathcal{E}} \ F_1$, and $r_0 \ \widehat{\mathcal{E}} \ r_1$. We show that clause 2a holds by case analysis on the different redexes.

- If $r_0 = \langle v_0\rangle$ and $r_1 = \langle v_1\rangle$ with $v_0 \ \widehat{\mathcal{E}}^{\mathrm{v}} \ v_1$, then $p_0 \rightarrow_{\mathrm{v}} F_0[v_0]$ and $p_1 \rightarrow_{\mathrm{v}} F_1[v_1]$. We have $F_0[v_0] \ \widehat{\mathcal{E}} \ F_1[v_1]$, as wished.
- Suppose $r_0 = (\lambda x.t_0')v_0$ and $r_1 = (\lambda x.t_1')v_1$ with $\lambda x.t_0' \ \widehat{\mathcal{E}}^{\mathrm{v}} \ \lambda x.t_1'$ and $v_0 \ \widehat{\mathcal{E}}^{\mathrm{v}} \ v_1$. Then $p_0 \rightarrow_{\mathrm{v}} F_0[t_0'\{v_0/x\}]$ and $p_1 \rightarrow_{\mathrm{v}} F_1[t_1'\{v_1/x\}]$. By Lemma 38 and because $\mathcal{X}$ is a congruence, we have $F_0[t_0'\{v_0/x\}] \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{v}}} \ F_1[t_1'\{v_1/x\}]$, as wished.
- If $r_0 = \langle E_0[\mathcal{S}k.t_0']\rangle$ and $r_1 = \langle E_1[\mathcal{S}k.t_1']\rangle$ with $E_0 \ \widetilde{\mathcal{E}} \ E_1$ and $t_0' \ \widehat{\mathcal{E}}^{\mathrm{v}} \ t_1'$. Then $p_0 \rightarrow_{\mathrm{v}} F_0[\langle t_0'\{\lambda x.\langle E_0[x]\rangle/k\}\rangle]$ and $p_1 \rightarrow_{\mathrm{v}} F_1[\langle t_1'\{\lambda x.\langle E_1[x]\rangle/k\}\rangle]$. From $E_0 \ \widetilde{\mathcal{E}} \ E_1$, we deduce $\lambda x.\langle E_0[x]\rangle \ \widehat{\mathcal{E}}^{\mathrm{v}} \ \lambda x.\langle E_1[x]\rangle$, so by Lemma 24, we have $F_0[\langle t_0'\{\lambda x.\langle E_0[x]\rangle/k\}\rangle] \ \widehat{\mathcal{E}} \ F_1[\langle t_1'\{\lambda x.\langle E_1[x]\rangle/k\}\rangle]$, as wished.

We now prove the bisimulation property (up to environment) for elements in $\mathcal{X}_1$. Let $F_0[t_0] \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{v}}} \ F_1[t_1]$, so that $t_0 \ \mathcal{Y}_{\mathcal{E}} \ t_1$ and $F_0 \ \widetilde{\mathcal{E}} \ F_1$. If $F_0[t_0]$ and $F_1[t_1]$ are not both programs, then for all $E_0 \ \widetilde{\mathcal{E}} \ E_1$, we have $\langle E_0[F_0[t_0]]\rangle \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{v}}} \ \langle E_1[F_1[t_1]]\rangle$, hence clause 1 holds. Suppose $F_0[t_0]$, $F_1[t_1]$ are programs p_0, p_1. We distinguish two cases. First, suppose t_0 and t_1 are programs. If $t_0 \rightarrow_{\mathrm{v}} p_0'$, then $p_0 \rightarrow_{\mathrm{v}} F_0[p_0']$. Because $t_0 \ \mathcal{Y}_{\mathcal{E}} \ t_1$, there exists p_1' such that $t_1 \rightarrow_{\mathrm{v}}^* p_1'$ and $p_0' \ \mathcal{Y}_{\mathcal{E}} \ p_1'$. We have $F_0[p_0'] \ \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{v}}} \ F_1[p_1']$ and $p_1 \rightarrow_{\mathrm{v}}^* F_1[p_1']$, therefore clause 2a holds. If $t_0 \rightarrow_{\mathrm{v}} v_0$, then $p_0 \rightarrow_{\mathrm{v}} F_0[v_0]$. Because $t_0 \ \mathcal{Y}_{\mathcal{E}} \ t_1$, there exists v_1 such that $t_1 \rightarrow_{\mathrm{v}}^* v_1$ and $\mathcal{E}' = \{(v_0, v_1)\} \cup \mathcal{E} \in \mathcal{Y}$. Hence $p_1 \rightarrow_{\mathrm{v}}^* F_1[v_1]$, and we have $F_0[v_0] \ \mathcal{X}_{\widehat{\mathcal{E}'}^{\mathrm{v}}} \ F_1[v_1]$, therefore clause 2a holds (up to environment).

In the second case, t_0 and t_1 are not both programs. Then we can write $p_0 = F_0'[\langle E_0[t_0]\rangle]$ and $p_1 = F_1'[\langle E_1[t_1]\rangle]$ for some $F_0' \ \widetilde{\mathcal{E}} \ F_1'$ and $E_0 \ \widetilde{\mathcal{E}} \ E_1$. Because $t_0 \ \mathcal{Y}_{\mathcal{E}} \ t_1$ and since $\mathcal{Y}$ is an environmental bisimulation, we have $\langle E_0[t_0]\rangle \ \mathcal{Y}_{\mathcal{E}}$

$\langle E_1[t_1]\rangle$. If $\langle E_0[t_0]\rangle \rightarrow_{\mathrm{v}} p'_0$, then there exists p'_1 such that $\langle E_1[t_1]\rangle \rightarrow^*_{\mathrm{v}} p'_1$ and $p'_0 \; \mathcal{Y}_{\mathcal{E}} \; p'_1$. Therefore, $p_0 \rightarrow_{\mathrm{v}} F_0{}'[p'_0]$, $p_1 \rightarrow_{\mathrm{v}} F_1{}'[p'_1]$, and $F_0{}'[p'_0] \; \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{v}}} \; F_1{}'[p'_1]$, hence clause 2a holds. If $\langle E_0[t_0]\rangle \rightarrow_{\mathrm{v}} v_0$, then there exists v_1 such that $\langle E_1[t_1]\rangle \rightarrow^*_{\mathrm{v}} v_1$ and $\mathcal{E}' = \{(v_0, v_1)\} \cup \mathcal{E} \subseteq \mathcal{Y}$. Therefore $p_0 \rightarrow_{\mathrm{v}} F_0{}'[v_0]$, $p_1 \rightarrow_{\mathrm{v}} F_1{}'[v_1]$, and $F_0{}'[v_0] \; \mathcal{X}_{\widehat{\mathcal{E}'}^{\mathrm{v}}} \; F_1{}'[v_1]$, hence clause 2a holds (up to environment).

We finally prove the clause 3 of the bisimulation. The only environments in $\mathcal{X}$ are of the form $\widehat{\mathcal{E}}^{\mathrm{v}}$. Let $\lambda x.t_0 \; \widehat{\mathcal{E}}^{\mathrm{v}} \; \lambda x.t_1$ and $v_0 \; \widehat{\mathcal{E}}^{\mathrm{v}} \; v_1$. By Lemma 38, we have $t_0\{v_0/x\} \; \mathcal{X}_{\widehat{\mathcal{E}}^{\mathrm{v}}} \; t_1\{v_1/x\}$, hence the result holds.

Lemma 39. *If $\lambda x.t_0 \mathrel{\dot{\simeq}} \lambda x.t_1$, then $t_0\{v/x\} \mathrel{\dot{\simeq}} t_1\{v/x\}$.*

Proof. If $\lambda x.t_0 \mathrel{\dot{\simeq}} \lambda x.t_1$, then $\langle \lambda x.t_0\rangle \mathrel{\dot{\simeq}} \langle \lambda x.t_1\rangle$ by clause 1. Since $\langle \lambda x.t_0\rangle \rightarrow_{\mathrm{v}} \lambda x.t_0$ and $\langle \lambda x.t_1\rangle \rightarrow_{\mathrm{v}} \lambda.t_1$, we have $\{(\lambda x.t_0, \lambda x.t_1)\} \in \mathrel{\dot{\simeq}}$ by clause 2b. Let $\mathcal{E} = \{(\lambda x.t_0, \lambda x.t_1)\}$. By clause 3, for all v, we have $t_0\{v/x\} \approx_{\mathcal{E}} t_1\{v/x\}$, therefore $t_0\{v/x\} \mathrel{\dot{\simeq}} t_1\{v/x\}$ holds by weakening (Lemma 5).

Lemma 40. *If $\lambda x.t_0 \mathrel{\widehat{\dot{\simeq}}} \lambda x.t \mathrel{\dot{\simeq}} \lambda x.t_1$ and $v_0 \mathrel{\widehat{\dot{\simeq}}} v \mathrel{\dot{\simeq}} v_1$ then $t_0\{v_0/x\} \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} t_1\{v_1/x\}$.*

Proof. We proceed by case analysis on $\lambda x.t_0 \mathrel{\widehat{\dot{\simeq}}} \lambda x.t$.

Suppose $\lambda x.t_0 \mathrel{\dot{\simeq}} \lambda x.t$. We have $t_0\{v_0/x\} \mathrel{\widehat{\dot{\simeq}}} t_0\{v/x\}$ by Lemma 24, $t_0\{v/x\} \mathrel{\dot{\simeq}} t\{v/x\}$ by Lemma 39, $t\{v/x\} \mathrel{\dot{\simeq}} t\{v_1/x\}$ by Lemma 36, and $t\{v_1/x\} \mathrel{\dot{\simeq}} t_1\{v_1/x\}$ by Lemma 39. Finally, $t_0\{v_0/x\} \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} t_1\{v_1/x\}$ holds using transitivity of $\dot{\simeq}$.

Suppose $t_0 \mathrel{\widehat{\dot{\simeq}}} t$ with $\mathsf{fv}(t_0) \cup \mathsf{fv}(t) \subseteq \{x\}$. We have $t_0\{v_0/x\} \mathrel{\widehat{\dot{\simeq}}} t\{v/x\}$ by Lemma 24, $t\{v/x\} \mathrel{\dot{\simeq}} t\{v_1/x\}$ by Lemma 36, and $t\{v_1/x\} \mathrel{\dot{\simeq}} t_1\{v_1/x\}$ by Lemma 39. Finally, $t_0\{v_0/x\} \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} t_1\{v_1/x\}$ holds using transitivity of $\dot{\simeq}$.

Programs are either value programs or can be decomposed in contexts F, E, and a redex r. We extend this result to related programs $p_0 \mathrel{\widehat{\dot{\simeq}}} p_1$, and see how they can be decomposed.

Lemma 41. *If $p_0 \mathrel{\widehat{\dot{\simeq}}} p_1$ then we have one of the following cases:*

- $p_0 \mathrel{\dot{\simeq}} p_1$*;*
- $p_0 = \langle v_0\rangle$*;*
- $p_0 = F_0[\langle E_0[t_0]\rangle]$, $p_1 = F_1[\langle E_1[t_1]\rangle]$, $F_0 \mathrel{\widetilde{\dot{\simeq}}} F_1$, $E_0 \mathrel{\widetilde{\dot{\simeq}}} E_1$, $t_0 \mathrel{\dot{\simeq}} t_1$ *and* $t_0 \rightarrow_{\mathrm{v}} t'_0$ *or* t_0 *is stuck;*
- $p_0 = F_0[\langle E_0[r_0]\rangle]$, $p_1 = F_1[\langle E_1[t_1]\rangle]$, $F_0 \mathrel{\widetilde{\dot{\simeq}}} F_1$, $E_0 \mathrel{\widetilde{\dot{\simeq}}} E_1$, $r_0 \mathrel{\widehat{\dot{\simeq}}} t_1$ *but* $r_0 \mathrel{\dot{\not\simeq}} t_1$*.*

Proof. We prove a more general result on $t_0 \mathrel{\widehat{\dot{\simeq}}} t_1$. We have either

- $t_0 \mathrel{\dot{\simeq}} t_1$;
- $t_0 = v_0$;

- $t_0 = E_0[t'_0]$, $t_1 = E_1[t'_1]$, $E_0 \mathrel{\widetilde{\dot{\simeq}}} E_1$, $t'_0 \mathrel{\dot{\simeq}} t'_1$, and $t'_0 \rightarrow_{\mathrm{v}} t''_0$ or t_0 is stuck;
- $t_0 = F_0[\langle E_0[t'_0]\rangle]$, $t_1 = F_1[\langle E_1[t'_1]\rangle]$, $F_0 \mathrel{\widetilde{\dot{\simeq}}} F_1$, $E_0 \mathrel{\widetilde{\dot{\simeq}}} E_1$, $t'_0 \mathrel{\dot{\simeq}} t'_1$, and $t'_0 \rightarrow_{\mathrm{v}} t''_0$ or t'_0 is stuck;
- $t_0 = E_0[r_0]$, $t_1 = E_1[t'_1]$, $E_0 \mathrel{\widetilde{\dot{\simeq}}} E_1$, $r_0 \mathrel{\widehat{\dot{\simeq}}} t'_1$ but $r_0 \mathrel{\dot{\not\simeq}} t'_1$.
- $t_0 = F_0[\langle E_0[r_0]\rangle]$, $t_1 = F_1[\langle E_1[t'_1]\rangle]$, $F_0 \mathrel{\widetilde{\dot{\simeq}}} F_1$, $E_0 \mathrel{\widetilde{\dot{\simeq}}} E_1$, $r_0 \mathrel{\widehat{\dot{\simeq}}} t'_1$ but $r_0 \mathrel{\dot{\not\simeq}} t'_1$.

The proof is easy by induction on $t_0 \mathrel{\widehat{\dot{\simeq}}} t_1$ but tedious.

Lemma 42. *If $v_0 \mathrel{\widehat{\dot{\simeq}}} t_1$, then there exists v_1 such that $\langle t_1\rangle \rightarrow_{\mathrm{v}}^* v_1$ and $v_0 \mathrel{\widehat{\dot{\simeq}}} v_1$.*

Proof. We have two cases to consider. If $v_0 \mathrel{\dot{\simeq}} t_1$, then $\langle v_0\rangle \mathrel{\dot{\simeq}} \langle v_1\rangle$, and because $\langle v_0\rangle \rightarrow_{\mathrm{v}} v_0$, there exists v_1 such that $\langle t_1\rangle \rightarrow_{\mathrm{v}}^* v_1$ and $\{(v_0, v_1)\} \in \dot{\approx}$. Because $\rightarrow_{\mathrm{v}}^* \subseteq \dot{\simeq}$, we have $v_0 \mathrel{\dot{\simeq}} v_1$, as wished. Otherwise, t_1 is a value v_1, and the result holds trivially.

Lemma 43. *Let $t_0 \mathrel{\dot{\simeq}} t_1$ so that $t_0 \rightarrow_{\mathrm{v}} t'_0$ or t_0 is stuck, and $E_0 \mathrel{\widetilde{\dot{\simeq}}} E_1$. There exist p'_0, p'_1 such that $\langle E_0[t_0]\rangle \rightarrow_{\mathrm{v}} p'_0$, $\langle E_1[t_1]\rangle \rightarrow_{\mathrm{v}}^* p'_1$, and $p'_0 \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} p'_1$.*

Proof. Suppose t_0 and t_1 are both programs. Then t_0 cannot be stuck, and we have $t_0 \rightarrow_{\mathrm{v}} t'_0$. By bisimilarity, there exists p'_1 such that $t_1 \rightarrow_{\mathrm{v}}^* p'_1$ and $t'_0 \mathrel{\dot{\simeq}} p'_1$. We have $\langle E_0[t_0]\rangle \rightarrow_{\mathrm{v}} \langle E_0[t'_0]\rangle$, $\langle E_1[t_1]\rangle \rightarrow_{\mathrm{v}}^* \langle E_1[p'_1]\rangle$, and $\langle E_0[t'_0]\rangle \mathrel{\widehat{\dot{\simeq}}} \langle E_1[p'_1]\rangle$, hence the result holds.

Suppose t_0 and t_1 are not both programs. Because $t_0 \mathrel{\dot{\simeq}} t_1$, we have $\langle E_1[t_0]\rangle \mathrel{\dot{\simeq}} \langle E_1[t_1]\rangle$. From $t_0 \rightarrow_{\mathrm{v}} t'_0$ or t_0 is stuck, we know that $\langle E_1[t_0]\rangle \rightarrow_{\mathrm{v}} p''_0$ for some p''_0 and $\langle E_0[t_0]\rangle \rightarrow_{\mathrm{v}} p'_0$ with $p'_0 \mathrel{\widehat{\dot{\simeq}}} p''_0$. By bisimilarity, there exists p'_1 such that $\langle E_1[t_1]\rangle \rightarrow_{\mathrm{v}}^* p'_1$ and $p''_0 \mathrel{\dot{\simeq}} p'_1$. We have $p'_0 \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} p'_1$, hence the result holds.

Lemma 44. *Let $\lambda x.t_0 \mathrel{\widehat{\dot{\simeq}}} t_1^1$, $v_0 \mathrel{\widehat{\dot{\simeq}}} t_1^2$, and $E_0 \mathrel{\widetilde{\dot{\simeq}}} E_1$. There exist p_0, p_1 such that $\langle E_0[(\lambda x.t_0)\, v_0]\rangle \rightarrow_{\mathrm{v}} p_0$, $\langle E_1[t_1^1\, t_1^2]\rangle \rightarrow_{\mathrm{v}}^* p_1$, and $p_0 \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} p_1$.*

Proof. We have four different cases:

- Suppose $\lambda x.t_0 \mathrel{\widehat{\dot{\simeq}}} \lambda x.t_1$, $v_0 \mathrel{\widehat{\dot{\simeq}}} v_1$. By Lemma 40, we have $t_0\{v_0/x\} \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} t_1\{v_1/x\}$. We have $\langle E_0[(\lambda x.t_0)\, v_0]\rangle \rightarrow_{\mathrm{v}} \langle E_0[t_0\{t_0/x\}]\rangle$, $\langle E_1[(\lambda x.t_1)\, v_1]\rangle \rightarrow_{\mathrm{v}} \langle E_1[t_1\{t_1/x\}]\rangle$, and we have $\langle E_0[t_0\{t_0/x\}]\rangle \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} \langle E_1[t_1\{t_1/x\}]\rangle$ by congruence of $\widehat{\dot{\simeq}}$ and Lemma 37. Therefore, we have the required result.
- Suppose $\lambda x.t_0 \mathrel{\widehat{\dot{\simeq}}} \lambda x.t_1$ with $t_0 \mathrel{\widehat{\dot{\simeq}}} t_1$ and $v_0 \mathrel{\dot{\simeq}} t_1^2$. By bisimilarity, we have $\langle E_1[(\lambda x.t_1)\, v_0]\rangle \mathrel{\dot{\simeq}} \langle E_1[(\lambda x.t_1)\, t_1^2]\rangle$, therefore there exists p_1 such that $\langle E_1[(\lambda x.t_1)\, t_1^2]\rangle \rightarrow_{\mathrm{v}}^* p_1$ and $\langle E_1[t_1\{v_0/x\}]\rangle \mathrel{\dot{\simeq}} p_1$. We get $\langle E_0[(\lambda x.t_0)\, v_0]\rangle \rightarrow_{\mathrm{v}} \langle E_0[t_0\{v_0/x\}]\rangle$ and $\langle E_0[t_0\{v_0/x\}]\rangle \mathrel{\widehat{\dot{\simeq}}} \langle E_1[t_1\{v_0/x\}]\rangle$, and hence the result holds.
- Suppose $\lambda x.t_0 \mathrel{\dot{\simeq}} t_1^1$ and $v_0 \mathrel{\widehat{\dot{\simeq}}} v_1$. By bisimilarity, we have $\langle E_1[(\lambda x.t_0)\, v_0]\rangle \mathrel{\dot{\simeq}} \langle E_1[t_1^1\, v_1]\rangle$, therefore there exists a program p_1 such that $\langle E_1[t_1^1\, v_1]\rangle \rightarrow_{\mathrm{v}}^* p_1$ and $\langle E_1[t_0\{v_0/x\}]\rangle \mathrel{\dot{\simeq}} p_1$. We have $\langle E_0[(\lambda x.t_0)\, v_0]\rangle \rightarrow_{\mathrm{v}} \langle E_0[t_0\{v_0/x\}]\rangle$ and $\langle E_0[t_0\{v_0/x\}]\rangle \mathrel{\widehat{\dot{\simeq}}} \langle E_1[t_0\{v_0/x\}]\rangle$, hence the result holds.

– Suppose $\lambda x.t_0 \mathrel{\dot{\simeq}} t_1^1$ and $v_0 \mathrel{\dot{\simeq}} t_1^2$. By bisimilarity, we have $\langle E_1[(\lambda x.t_0)\, v_0]\rangle \mathrel{\dot{\simeq}} \langle E_1[t_1^1\, v_0]\rangle$ and $\langle E_1[t_1^1\, v_0]\rangle \mathrel{\dot{\simeq}} \langle E_1[t_1^1\, t_1^2]\rangle$, therefore there exists p_1 such that $\langle E_1[t_1^1\, t_1^2]\rangle \rightarrow_{\mathrm{v}}^{*} p_1$ and $\langle E_1[t_0\{v_0/x\}]\rangle \mathrel{\dot{\simeq}} p_1$. We get $\langle E_0[(\lambda x.t_0)\, v_0]\rangle \rightarrow_{\mathrm{v}} \langle E_0[t_0\{v_0/x\}]\rangle$ and $\langle E_0[t_0\{v_0/x\}]\rangle \mathrel{\widehat{\dot{\simeq}}} \langle E_1[t_0\{v_0/x\}]\rangle$, and hence the result holds.

Lemma 45. *If $E[\mathcal{S}k.t_0] \mathrel{\widehat{\dot{\simeq}}} t_1$ and $E_0 \mathrel{\widetilde{\dot{\simeq}}} E_1$, then there exist p_0, p_1 such that $\langle E_0[E[\mathcal{S}k.t_0]]\rangle \rightarrow_{\mathrm{v}} p_0$, $\langle E_1[t_1]\rangle \rightarrow_{\mathrm{v}}^{*} p_1$, and $p_0 \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} p_1$.*

Proof. We proceed by induction on $E[\mathcal{S}k.t_0] \mathrel{\widehat{\dot{\simeq}}} t_1$.

If $E[\mathcal{S}k.t_0] \mathrel{\dot{\simeq}} t_1$, then by bisimilarity we have $\langle E_1[E[\mathcal{S}k.t_0]]\rangle \mathrel{\dot{\simeq}} \langle E_1[t_1]\rangle$. Because $\langle E_1[E[\mathcal{S}k.t_0]]\rangle \rightarrow_{\mathrm{v}} p_0'$, there exists p_1 such that $\langle E_1[t_1]\rangle \rightarrow_{\mathrm{v}}^{*} p_1$ and $p_0' \mathrel{\dot{\simeq}} p_1$. We also have $\langle E_0[E[\mathcal{S}k.t_0]]\rangle \rightarrow_{\mathrm{v}} p_0$ with $p_0 \mathrel{\widehat{\dot{\simeq}}} p_0'$, hence the result holds.

If $E = \square$, $t_1 = \mathcal{S}k.t_1'$ with $t_0 \mathrel{\widehat{\dot{\simeq}}} t_1'$, then we have $\langle t_0\{\lambda k.\langle E_0[x]\rangle/k\}\rangle \mathrel{\widehat{\dot{\simeq}}} \langle t_1'\{\lambda k.\langle E_1[x]\rangle/k\}\rangle$ by Lemma 24, and because we have the two reductions $\langle E_0[\mathcal{S}k.t_0]\rangle \rightarrow_{\mathrm{v}} \langle t_0\{\lambda k.\langle E_0[x]\rangle/k\}\rangle$ and $\langle E_1[\mathcal{S}k.t_1]\rangle \rightarrow_{\mathrm{v}} \langle t_1'\{\lambda k.\langle E_1[x]\rangle/k\}\rangle$, we obtain the required result.

Suppose $E[\mathcal{S}k.t_0] = v_0\, E'[\mathcal{S}k.t_0]$, $t_1 = t_1^1\, t_1^2$ with $v_0 \mathrel{\widehat{\dot{\simeq}}} t_1^1$ and $E'[\mathcal{S}k.t_0] \mathrel{\widehat{\dot{\simeq}}} t_1^2$. We distinguish two cases. If $t_1^1 = v_1$, then by the induction hypothesis, there exist p_0, p_1 such that $\langle E_0[v_0\, E'[\mathcal{S}k.t_0]]\rangle \rightarrow_{\mathrm{v}} p_0$, $\langle E_1[v_1\, t_1^2]\rangle \rightarrow_{\mathrm{v}}^{*} p_1$, and $p_0 \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} p_1$, therefore we have the required result. Suppose now $v_0 \mathrel{\dot{\simeq}} t_1^1$. By the induction hypothesis, there exist p_0, p_0' such that $\langle E_0[v_0\, E'[\mathcal{S}k.t_0]]\rangle \rightarrow_{\mathrm{v}} p_0$, $\langle E_1[v_0\, t_1^2]\rangle \rightarrow_{\mathrm{v}}^{*} p_0'$, and $p_0 \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} p_0'$. From $v_0 \mathrel{\dot{\simeq}} t_1^1$, we know that $\langle E_1[v_0\, t_1^2]\rangle \mathrel{\dot{\simeq}} \langle E_1[t_1^1\, t_1^2]\rangle$. By bisimilarity, there exists p_1 such that $\langle E_1[t_1^1\, t_1^2]\rangle \rightarrow_{\mathrm{v}}^{*} p_1$ and $p_0' \mathrel{\dot{\simeq}} p_1$. Therefore we have $p_0 \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} p_1$, hence the result holds.

Suppose $E[\mathcal{S}k.t_0] = E'[\mathcal{S}k.t_0]\, t$, $t_1 = t_1^1\, t_1^2$ with $E'[\mathcal{S}k.t_0] \mathrel{\widehat{\dot{\simeq}}} t_1^1$ and $t \mathrel{\widehat{\dot{\simeq}}} t_1^2$. By the induction hypothesis, there exist p_0, p_1 such that $\langle E_0[E'[\mathcal{S}k.t_0]\, t]\rangle \rightarrow_{\mathrm{v}} p_0$, $\langle E_1[t_1^1\, t_1^2]\rangle \rightarrow_{\mathrm{v}}^{*} p_1$, and $p_0 \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} p_1$, therefore we have the required result.

Lemma 46. *$t_0 \mathrel{\dot{\simeq}} t_1$ implies $C[t_0] \approx_{\widehat{\dot{\simeq}}}^{\mathrm{nf}} C[t_1]$.*

Proof. We prove that

$$\mathcal{X} = \{(\widehat{\dot{\simeq}}^{\mathrm{v}}, t_0, t_1) \mid t_0 \mathrel{\widehat{\dot{\simeq}}} t_1\} \cup \{\widehat{\dot{\simeq}}^{\mathrm{v}}\}$$

is a bisimulation up to bisimilarity. Note that by definition of $\mathcal{X}$, we have $t \mathrel{\mathcal{X}_{\widehat{\dot{\simeq}}^{\mathrm{v}}}} t'$ iff $t \mathrel{\widehat{\dot{\simeq}}} t'$.

Let $t_0 \mathrel{\mathcal{X}_{\widehat{\dot{\simeq}}^{\mathrm{v}}}} t_1$. Suppose t_0 and t_1 are not both programs. Then for all $E_0 \mathrel{\widetilde{\dot{\simeq}}} E_1$, we have $\langle E_0[t_0]\rangle \mathrel{\mathcal{X}_{\widehat{\dot{\simeq}}^{\mathrm{v}}}} \langle E_1[t_1]\rangle$, hence clause 1 is satisfied.

Suppose t_0 and t_1 are both programs p_0, p_1. By Lemma 41, we have several possibilities. If $p_0 \mathrel{\dot{\simeq}} p_1$, then the result holds trivially. Suppose $p_0 = \langle v_0\rangle$, $p_1 = \langle t_1'\rangle$ with $v_0 \mathrel{\widehat{\dot{\simeq}}} t_1'$. By Lemma 42, there exists v_1 such that $\langle t_1'\rangle \rightarrow_{\mathrm{v}}^{*} v_1$ and $v_0 \mathrel{\widehat{\dot{\simeq}}} v_1$. We have $\{(v_0, v_1)\} \cup \widehat{\dot{\simeq}}^{\mathrm{v}} = \widehat{\dot{\simeq}}^{\mathrm{v}} \in \mathcal{X}$, hence the result holds.

Suppose $p_0 = F_0[\langle E_0[t'_0]\rangle]$ and $p_1 = F_1[\langle E_1[t'_1]\rangle]$, with $F_0 \mathrel{\widetilde{\dot{\simeq}}} F_1$, $E_0 \mathrel{\widetilde{\dot{\simeq}}} E_1$, $t'_0 \mathrel{\dot{\simeq}} t'_1$, and $t_0 \rightarrow_{\mathrm{v}} t'_0$ or t_0 is stuck. By Lemma 43, there exist p'_0, p'_1 such that $\langle E_0[t'_0]\rangle \rightarrow_{\mathrm{v}} p'_0$, $\langle E_1[t'_1]\rangle \rightarrow^*_{\mathrm{v}} p'_1$, and $p'_0 \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} p'_1$. By definition of $\widehat{\dot{\simeq}}$ and Lemma 37, we have $F_0[p'_0] \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} F_1[p'_1]$. Moreover $p_0 \rightarrow_{\mathrm{v}} F_0[p'_0]$ and $p_1 \rightarrow^*_{\mathrm{v}} F_1[p'_1]$, hence the result holds.

The last possibility is $p_0 = F_0[\langle E_0[r_0]\rangle]$, $p_1 = F_1[\langle E_1[t'_1]\rangle]$, with $F_0 \mathrel{\widetilde{\dot{\simeq}}} F_1$, $E_0 \mathrel{\widetilde{\dot{\simeq}}} E_1$, $r_0 \mathrel{\widehat{\dot{\simeq}}} t'_1$, and $r_0 \mathrel{\dot{\not\simeq}} t'_1$. We discuss the three possible redexes. If $r_0 = \langle v_0\rangle$, then $t'_1 = \langle t''_1\rangle$ with $v_0 \mathrel{\widehat{\dot{\simeq}}} t''_1$. By Lemma 42, there exists v_1 such that $\langle t''_1\rangle \rightarrow^*_{\mathrm{v}} v_1$ and $v_0 \mathrel{\widehat{\dot{\simeq}}} v_1$. Then we have $p_0 \rightarrow_{\mathrm{v}} F_0[\langle E_0[v_0]\rangle]$ and $p_1 \rightarrow^*_{\mathrm{v}} F_1[\langle E_1[v_1]\rangle]$ with $F_0[\langle E_0[v_0]\rangle] \mathrel{\widehat{\dot{\simeq}}} F_1[\langle E_1[v_1]\rangle]$, hence the result holds. If $r_0 = v_0^1\, v_0^2$, then $t'_1 = t_1^1\, t_1^2$ with $v_0^1 \mathrel{\widehat{\dot{\simeq}}} t_1^1$, and $v_0^2 \mathrel{\widehat{\dot{\simeq}}} t_1^2$. By Lemma 44, there exist p'_0, p'_1 such that $E_0[r_0] \rightarrow_{\mathrm{v}} p'_0$, $E_1[t'_1] \rightarrow^*_{\mathrm{v}} p'_1$, and $p'_0 \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} p'_1$. Therefore we have $p_0 \rightarrow_{\mathrm{v}} F_0[p'_0]$, $p_1 \rightarrow^*_{\mathrm{v}} F_1[p'_1]$, with $F_0[p'_0] \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} F_1[p'_1]$ (by Lemma 37 and definition of $\widehat{\dot{\simeq}}$), hence the result holds. The last case is $r_0 = \langle E[\mathcal{S}k.t'_0]\rangle$; then $t'_1 = \langle t''_1\rangle$ with $E[\mathcal{S}k.t'_0] \mathrel{\widehat{\dot{\simeq}}} t''_1$. By Lemma 45, there exist p'_0, p'_1 such that $r_0 \rightarrow_{\mathrm{v}} p'_0$, $t'_1 \rightarrow^*_{\mathrm{v}} p'_1$ and $p'_0 \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} p'_1$. Therefore we have $p_0 \rightarrow_{\mathrm{v}} F_0[\langle E_0[p'_0]\rangle]$, $p_1 \rightarrow^*_{\mathrm{v}} F_1[\langle E_1[p'_1]\rangle]$, with $F_0[\langle E_0[p'_0]\rangle] \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} F_1[\langle E_1[p'_1]\rangle]$ (by Lemma 37 and definition of $\widehat{\dot{\simeq}}$), hence the result holds.

Finally, let $\lambda x.t_0 \mathrel{\widehat{\dot{\simeq}}} \lambda x.t_1$ and $v_0 \mathrel{\widehat{\dot{\simeq}}} v_1$. By Lemma 40, we get $t_0\{v_0/x\} \mathrel{\widehat{\dot{\simeq}}\dot{\simeq}} t_1\{v_1/x\}$, hence the required result holds.

Lemma 47. *We have* $\ddot{\approx}_c \subseteq \dot{\simeq}$.

Proof. We prove that $\mathcal{X} = \{(\ddot{\approx}_c{}^{\mathrm{v}}, t_0, t_1) \mid t_0 \mathrel{\ddot{\approx}_c} t_1\} \cup \{\approx_c{}^{\mathrm{v}}\}$ is a big-step environmental bisimulation for programs.

Let $t_0 \mathrel{\mathcal{X}_{\ddot{\approx}_c{}^{\mathrm{v}}}} t_1$ such that t_0 and t_1 are not both programs. Because $\ddot{\approx}_c$ is a congruence w.r.t. evaluation contexts, we have $\langle E_0[t_0]\rangle \mathrel{\ddot{\approx}_c} \langle E_1[t_1]\rangle$ for all $E_0 \mathrel{\widetilde{\ddot{\approx}_c{}^{\mathrm{v}}}} E_1$, i.e., $\langle E_0[t_0]\rangle \mathrel{\mathcal{X}_{\ddot{\approx}_c{}^{\mathrm{v}}}} \langle E_1[t_1]\rangle$ as wished.

Let $p_0 \mathrel{\mathcal{X}_{\ddot{\approx}_c{}^{\mathrm{v}}}} p_1$ such that $p_0 \rightarrow^*_{\mathrm{v}} v_0$. We have $\langle p_0\rangle \rightarrow^*_{\mathrm{v}} v_0$, so by definition of $\ddot{\approx}_c$, there exists v_1 such that $\langle p_1\rangle \rightarrow^*_{\mathrm{v}} v_1$, which implies $p_1 \rightarrow^*_{\mathrm{v}} v_1$. By Lemma 35, we have $p_0 \mathrel{\dot{\simeq}} v_0$ and $p_1 \mathrel{\dot{\simeq}} v_1$, which implies $p_0 \mathrel{\ddot{\approx}_c} v_0$ and $p_1 \mathrel{\ddot{\approx}_c} v_1$ by soundness of $\dot{\simeq}$. Transitivity of $\ddot{\approx}_c$ gives $v_0 \mathrel{\ddot{\approx}_c} v_1$, hence we have $\ddot{\approx}_c{}^{\mathrm{v}} \cup \{(v_0, v_1)\} = \ddot{\approx}_c{}^{\mathrm{v}} \in \mathcal{X}$, as wished.

Let $\lambda x.t_0 \mathrel{\ddot{\approx}_c} \lambda x.t_1$ and $v_0 \mathrel{\widehat{\ddot{\approx}_c{}^{\mathrm{v}}}} v_1$. By congruence of $\ddot{\approx}_c$, we have $v_0 \mathrel{\ddot{\approx}_c} v_1$, and also $(\lambda x.t_0)\, v_0 \mathrel{\ddot{\approx}_c} (\lambda x.t_1)\, v_1$. Because $(\lambda x.t_0) \rightarrow_{\mathrm{v}} t_0\{v_0/x\}$, $(\lambda x.t_1)\, v_1 \rightarrow_{\mathrm{v}} t_1\{v_1/x\}$ and $\rightarrow_{\mathrm{v}} \subseteq \dot{\simeq} \subseteq \ddot{\approx}_c$, we have $t_0\{v_0/x\} \mathrel{\ddot{\approx}_c} t_1\{v_1/x\}$, i.e., $t_0\{v_0/x\} \mathrel{\mathcal{X}_{\ddot{\approx}_c{}^{\mathrm{v}}}} t_1\{v_1/x\}$, as wished.

C Bisimulation proofs

C.1 Proof of the $\mathcal{S}k.k\ t =_{\textsc{kh}} t$ Axiom

Let

$$\mathcal{E}_1 = \{(\lambda \overrightarrow{x}.t\sigma^0_1 \ldots \sigma^0_n, \lambda \overrightarrow{x}.\mathcal{S}k.k\ t\sigma^1_1 \ldots \sigma^1_n) \mid \\ \mathsf{fv}(t) \subseteq \overrightarrow{x} \cup \{x_1 \ldots x_n\}, k \notin \mathsf{fv}(t)\}$$

and

$$\mathcal{E}_2 = \{(v\sigma^0_1 \ldots \sigma^0_m, v\sigma^1_1 \ldots v\sigma^1_m) \mid \mathsf{fv}(v) \subseteq \{x_1 \ldots x_m\}\},$$

where σ^0_i and σ^1_i are of the form $\cdot\{\lambda \overrightarrow{y}.t_i/x_i\}$ and $\cdot\{\lambda \overrightarrow{y}.\mathcal{S}k.k\ t_i/x_i\}$ respectively for some t_i such that $k \notin \mathsf{fv}(t_i)$. Let $\mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_2$. We prove that

$$\mathcal{X} = \{(\mathcal{E}, v_1, v_2) \mid v_1\ \mathcal{E}_1\ v_2\} \\ \cup \{(\mathcal{E}, t\sigma^0_1 \ldots \sigma^0_m, \mathcal{S}k.k\ t\sigma^1_1 \ldots \sigma^1_m \mid \mathsf{fv}(t) \subseteq \{x_1 \ldots x_n\}, k \notin \mathsf{fv}(t)\} \\ \cup \{(\mathcal{E}, t\sigma^0_1 \ldots \sigma^0_n, t\sigma^1_1 \ldots \sigma^1_n), \mid \mathsf{fv}(t) \subseteq \{x_1 \ldots x_n\}\}$$

is a bisimulation for programs up to bisimilarity.

Let $v_1\ \mathcal{E}_1\ v_2$, and let $E_1\ \widetilde{\mathcal{E}}\ E_2$. Then we have $v_1 = \lambda \overrightarrow{x}.t\sigma^0_1 \ldots \sigma^0_n$, $v_2 = \lambda \overrightarrow{x}.\mathcal{S}k.k\ t\sigma^1_1 \ldots \sigma^1_n$, and E_1 and E_2 can be rewritten $E_1 = E\sigma^0_{n+1} \ldots \sigma^0_m$, and $E_2 = E\sigma^1_{n+1} \ldots \sigma^1_m$ for some E. We can rewrite $\langle E_1[v_1]\rangle$ and $\langle E_2[v_2]\rangle$ into

$$\langle E_1[v_1]\rangle = \langle E[y]\rangle\{\lambda \overrightarrow{x}.t/y\}\sigma^0_1 \ldots \sigma^0_m \\ \langle E_2[v_2]\rangle = \langle E[y]\rangle\{\lambda \overrightarrow{x}.\mathcal{S}k.k\ t/y\}\sigma^1_1 \ldots \sigma^1_m$$

for some fresh y. Hence, we have $\langle E_1[v_1]\rangle\ \mathcal{X}_{\mathcal{E}}\ \langle E_2[v_2]\rangle$, as wished.

Let $t_0 = t\sigma^0_1 \ldots \sigma^0_n\ \mathcal{X}_{\mathcal{E}}\ \mathcal{S}k.k\ t\sigma^1_1 \ldots \sigma^1_n = t_1$, and let $E_1\ \widetilde{\mathcal{E}}\ E_2$. Then E_1 and E_2 can be rewritten $E_1 = E\sigma^0_{n+1} \ldots \sigma^0_m$, and $E_2 = E\sigma^1_{n+1} \ldots \sigma^1_m$ for some E. Therefore

$$\langle E_1[t_0]\rangle = \langle E[t]\rangle\sigma^0_1 \ldots \sigma^0_m \\ \langle E_2[t_1]\rangle = \langle E[\mathcal{S}k.k\ t]\rangle\sigma^1_1 \ldots \sigma^1_m$$

but we have $\langle E[\mathcal{S}k.k\ t]\rangle\sigma^1_1 \ldots \sigma^1_m \doteq \langle E[t]\rangle\sigma^1_1 \ldots \sigma^1_m$ by the same reasoning as in Section 4.4, therefore we have $\langle E_1[t_0]\rangle\ \mathcal{X}_{\mathcal{E}}\!\doteq \langle E_2[t_1]\rangle$ as required.

Let $t_0 = t\sigma^0_1 \ldots \sigma^0_n\ \mathcal{X}_{\mathcal{E}}\ t\sigma^1_1 \ldots \sigma^1_n = t_1$. If t is not a program, then for all $E_1\ \widetilde{\mathcal{E}}\ E_2$, we can show that $E_1[t_0]\ \mathcal{X}_{\mathcal{E}}\ E_2[t_1]$ by rewriting E_1 and E_2 as in the previous cases. If t is a program p, then we distinguish several cases. If $p \rightarrow_{\mathrm{v}} p'$, then we conclude easily. If $p = \langle v\rangle$, then t_0 and t_1 reduce to values related by $\mathcal{E}_2$. If $p = \langle x_j\rangle$, then t_0 and t_1 reduce to values related by $\mathcal{E}_1$. If $p = \langle F[x_j\ v]\rangle$, then $t_0 = \langle F[(\lambda \overrightarrow{y}.t_j)\ v]\rangle\sigma^0_1 \ldots \sigma^0_n$ and $t_1 = \langle F[(\lambda \overrightarrow{y}.\mathcal{S}k.k\ t_j)\ v]\rangle\sigma^1_1 \ldots \sigma^1_n$. If $\overrightarrow{y} = y_0 \cup \overrightarrow{y'}$ with $\overrightarrow{y'}$ not empty, then $t_0 \rightarrow_{\mathrm{v}} \langle F[\lambda \overrightarrow{y'}.t_j\{v/y_0\}]\rangle\sigma^0_1 \ldots \sigma^0_n$ and $t_1 \rightarrow_{\mathrm{v}} \langle F[\lambda \overrightarrow{y'}.\mathcal{S}k.k\ t_j\{v/y_0\}]\rangle\sigma^1_1 \ldots \sigma^1_n$. The two resulting terms can be rewritten into

$p'\sigma_1^0 \ldots \sigma_j^0 \sigma_{j'}^0 \ldots \sigma_n^0$ and $p'\sigma_1^1 \ldots \sigma_j^1 \sigma_{j'}^1 \ldots \sigma_n^1$, where we have $p' = \langle F[x_{j'}]\rangle$, $\sigma_{j'}^0 = \cdot\{\lambda \overrightarrow{y'}.t_j\{v/y_0\}/x_{j'}\}$, $\sigma_{j'}^1 = \cdot\{\lambda \overrightarrow{y'}.\mathcal{S}k.k\, t_j\{v/y_0\}/x_{j'}\}$, and $x_{j'}$ is fresh. If $\overrightarrow{y} = y$, then $t_0 \to_{\mathrm{v}} \langle F[t_j\{v/y_0\}]\rangle \sigma_1^0 \ldots \sigma_n^0 = p_0$ and $t_1 \to_{\mathrm{v}} \langle F[\mathcal{S}k.k\, t_j\{v/y_0\}]\rangle \sigma_1^1 \ldots \sigma_n^1 = p_1$. By the same reasoning as in Section 4.4 (and with some case analysis on F), we have $\langle F[\mathcal{S}k.k\, t_j\{v/y_0\}]\rangle \sigma_1^1 \ldots \sigma_n^1 \doteq \langle F[t_j\{v/y_0\}]\rangle \sigma_1^1 \ldots \sigma_n^1$, therefore we have $p_0 \;\mathcal{X}_{\mathcal{E}} \dot{\simeq}\; p_1$, as wished.

Let $\lambda \overrightarrow{x}.t\sigma_1^0 \ldots \sigma_n^0 \;\mathcal{E}_1\; \lambda \overrightarrow{x}.\mathcal{S}k.k\, t\sigma_1^1 \ldots \sigma_n^1$ and $v_0 \;\widetilde{\mathcal{E}}\; v_1$. Then $v_0 = v\sigma_{n+1}^0 \ldots \sigma_m^0$ and $v_1 = v\sigma_{n+1}^1 \ldots \sigma_m^1$ for some v. If $\overrightarrow{x} = y \cup \overrightarrow{x'}$ with $\overrightarrow{x'}$ not empty, then we have $\lambda \overrightarrow{x'}.t\{v/y\}\sigma_1^0 \ldots \sigma_m^0 \;\mathcal{X}_{\mathcal{E}}\; \lambda \overrightarrow{x'}.\mathcal{S}k.k\, t\{v/y\}\sigma_1^1 \ldots \sigma_m^1$. If $\overrightarrow{x} = y$, then we also have $t\{v/y\}\sigma_1^0 \ldots \sigma_m^0 \;\mathcal{X}_{\mathcal{E}}\; \mathcal{S}k.k\, t\{v/y\}\sigma_1^1 \ldots \sigma_m^1$, therefore the result holds in both cases.

Let $\lambda x.t\sigma_1^0 \ldots \sigma_n^0 \;\mathcal{E}_2\; \lambda x.t\sigma_1^1 \ldots v\sigma_n^1$ and $v_0 \;\widetilde{\mathcal{E}}\; v_1$. Then $v_0 = v\sigma_{n+1}^0 \ldots \sigma_m^0$ and $v_1 = v\sigma_{n+1}^1 \ldots \sigma_m^1$ for some v. Then $t\{v/x\}\sigma_1^0 \ldots \sigma_m^0 \;\mathcal{X}_{\mathcal{E}}\; t\{v/x\}\sigma_1^1 \ldots \sigma_m^1$ holds.

C.2 Proof of the S-tail Axiom

Let t_0 and t_1 such that $k \notin \mathsf{fv}(t_1)$. We want to show that $(\lambda x.\mathcal{S}k.t_0)\; t_1 \doteq \mathcal{S}k.((\lambda x.t_0)\; t_1)$. To this end, we need to plug both terms in some context $\langle E\rangle$, and compare $\langle E[(\lambda x.\mathcal{S}k.t_0)\; t_1]\rangle$ with $\langle E[\mathcal{S}k.((\lambda x.t_0)\; t_1)]\rangle$. The second term reduces to $\langle(\lambda x.t_0\{\lambda y.\langle E[y]\rangle/k\})\; t_1\rangle$, so we in fact prove the following result.

Lemma 48. *We have* $\langle E[(\lambda x.\mathcal{S}k.t_0)\; t_1]\rangle \simeq \langle(\lambda x.t_0\{\lambda y.\langle E[y]\rangle/k\})\; t_1\rangle$.

Proof. To make the proof easier to follow, we introduce some notations. We write $\overrightarrow{\cdot}$ for a sequence of entities (e.g., $\overrightarrow{E}$ for a sequence of contexts). We write $\boldsymbol{E}$ for $E[(\lambda x.\mathcal{S}k.t_0)\,\square]$ and $\boldsymbol{E'}$ for $(\lambda x.t_0\{\lambda y.\langle E[y]\rangle/k\})\,\square$, so the problem becomes relating $\langle \boldsymbol{E}[t_1]\rangle$ and $\langle \boldsymbol{E'}[t_1]\rangle$.

Next, given a sequence $E_0 \ldots E_i$ of contexts such that $\mathsf{fv}(E_j) \subseteq \{k_0 \ldots k_{j-1}\}$ for all $0 \leq j \leq i$, we inductively define families of substitutions $\sigma_0^{\overrightarrow{E}} \ldots \sigma_i^{\overrightarrow{E}}$, $\delta_0^{\overrightarrow{E}} \ldots \delta_i^{\overrightarrow{E}}$ as follows:

$\sigma_0^{\overrightarrow{E}} = \cdot\{\lambda y.\langle \boldsymbol{E}[E_0[y]]\rangle/k_0\}$
$\delta_0^{\overrightarrow{E}} = \cdot\{\lambda y.\langle \boldsymbol{E'}[E_0[y]]\rangle/k_0\}$
$\sigma_j^{\overrightarrow{E}} = \cdot\{\lambda y.\langle \boldsymbol{E}[E_j\sigma_0^{\overrightarrow{E}} \ldots \sigma_{j-1}^{\overrightarrow{E}}[y]]\rangle/k_j\}$ if $j > 0$
$\delta_j^{\overrightarrow{E}} = \cdot\{\lambda y.\langle \boldsymbol{E'}[E_j\delta_0^{\overrightarrow{E}} \ldots \delta_{j-1}^{\overrightarrow{E}}[y]]\rangle/k_j\}$ if $j > 0$

Finally, given a term t and a sequence of contexts $F_0 \ldots F_i$, we inductively define families of terms $s_0^{t,\overrightarrow{F}} \ldots s_i^{t,\overrightarrow{F}}$, $u_0^{t,\overrightarrow{F}} \ldots u_i^{t,\overrightarrow{F}}$ as follows:

$s_0^{t,\overrightarrow{F}} = F_0[\langle \boldsymbol{E}[t]\rangle]$ $\qquad s_j^{t,\overrightarrow{F}} = F_j[\langle \boldsymbol{E}[s_{j-1}^{t,\overrightarrow{F}}]\rangle]$ if $j > 0$
$u_0^{t,\overrightarrow{F}} = F_0[\langle \boldsymbol{E'}[t]\rangle]$ $\qquad u_0^{t,\overrightarrow{F}} = F_j[\langle \boldsymbol{E'}[u_{j-1}^{t,\overrightarrow{F}}]\rangle]$ if $j > 0$

Note that the term we want to relate are $s_0^{t_1,\square}$ and $u_0^{t_1,\square}$. We let $\mathcal{E}$ ranges over environments of the form $\{(v\sigma_0^{\overrightarrow{E}} \ldots \sigma_i^{\overrightarrow{E}}, v\delta_0^{\overrightarrow{E}} \ldots \delta_i^{\overrightarrow{E}}) \mid \mathsf{fv}(v) \subseteq \{k_0 \ldots k_i\}\} \cup$

$\{(k_j\sigma_j^{\overrightarrow{E}}, k_j\delta_j^{\overrightarrow{E}})\}$. We prove that the relation

$$\mathcal{X} = \{(\mathcal{E}, t\sigma_0^{\overrightarrow{E}} \dots \sigma_i^{\overrightarrow{E}}, t\delta_0^{\overrightarrow{E}} \dots \delta_i^{\overrightarrow{E}}) \mid \mathsf{fv}(t) \subseteq \{k_0 \dots k_i\}\} \cup$$
$$\{(\mathcal{E}, \langle s_i^{t,\overrightarrow{F}}\rangle\sigma_0^{\overrightarrow{E'}} \dots \sigma_j^{\overrightarrow{E'}}, \langle u_i^{t,\overrightarrow{F}}\rangle\delta_0^{\overrightarrow{E'}} \dots \delta_j^{\overrightarrow{E'}}) \mid \mathsf{fv}(t) \cup \mathsf{fv}(\overrightarrow{F}) \subseteq \{k_0 \dots k_j\}\} \cup \{\mathcal{E}\}$$

is a bisimulation for programs. Let $t\sigma_0^{\overrightarrow{E}} \dots \sigma_i^{\overrightarrow{E}} \; \mathcal{X}_{\mathcal{E}} \; t\delta_0^{\overrightarrow{E}} \dots \delta_i^{\overrightarrow{E}}$ where t is not a program. Let $E_0 \; \widetilde{\mathcal{E}} \; E_1$; by definition of $\mathcal{E}$, we have $E_0 = E'\sigma_0^{\overrightarrow{E'}} \dots \sigma_j^{\overrightarrow{E''}}$ and $E_1 = E'\delta_0^{\overrightarrow{E}} \dots \delta_j^{\overrightarrow{E''}}$ for some E', $\overrightarrow{E''}$. With some renumbering and rewriting, we have $\langle E_0[t\sigma_0^{\overrightarrow{E}} \dots \sigma_i^{\overrightarrow{E}}]\rangle = \langle E'[t]\rangle\sigma_0^{\overrightarrow{E},\overrightarrow{E''}} \dots \sigma_{i+j+1}^{\overrightarrow{E},\overrightarrow{E''}}$ and $\langle E_1[t\sigma_0^{\overrightarrow{E}} \dots \sigma_i^{\overrightarrow{E}}]\rangle = \langle E'[t]\rangle\delta_0^{\overrightarrow{E},\overrightarrow{E''}} \dots \delta_{i+j+1}^{\overrightarrow{E},\overrightarrow{E''}}$: the two terms are in $\mathcal{X}$, as wished.

Let $\langle t\rangle\sigma_0^{\overrightarrow{E}} \dots \sigma_i^{\overrightarrow{E}} \; \mathcal{X}_{\mathcal{E}} \; \langle t\rangle\delta_0^{\overrightarrow{E}} \dots \delta_i^{\overrightarrow{E}}$. We have three cases for t.

If $\langle t\rangle \rightarrow_{\mathrm{v}} \langle t'\rangle$, we still have $\langle t'\rangle\sigma_0^{\overrightarrow{E}} \dots \sigma_i^{\overrightarrow{E}} \; \mathcal{X}_{\mathcal{E}} \; \langle t'\rangle\delta_0^{\overrightarrow{E}} \dots \delta_i^{\overrightarrow{E}}$. If $\langle t\rangle \rightarrow_{\mathrm{v}} v$ or $t = k_j$, then both terms reduce to values that are in $\mathcal{E}$, by definition of $\mathcal{E}$.

If $t = F[k_j\ v]$, then

$$\langle t\rangle\sigma_0^{\overrightarrow{E}} \dots \sigma_i^{\overrightarrow{E}} = F\sigma_0^{\overrightarrow{E}} \dots \sigma_i^{\overrightarrow{E}}[\lambda y.\langle \boldsymbol{E}[E_j\sigma_0^{\overrightarrow{E}} \dots \sigma_{j-1}^{\overrightarrow{E}}[y]]\rangle\ v\sigma_0^{\overrightarrow{E}} \dots \sigma_i^{\overrightarrow{E}}]$$
$$\langle t\rangle\delta_0^{\overrightarrow{E}} \dots \delta_i^{\overrightarrow{E}} = F\delta_0^{\overrightarrow{E}} \dots \delta_i^{\overrightarrow{E}}[\lambda y.\langle \boldsymbol{E'}[E_j\delta_0^{\overrightarrow{E}} \dots \delta_{j-1}^{\overrightarrow{E}}[y]]\rangle\ v\delta_0^{\overrightarrow{E}} \dots \delta_i^{\overrightarrow{E}}]$$

Reducing the β-redex in both terms, we obtain

$$\langle t\rangle\sigma_0^{\overrightarrow{E}} \dots \sigma_i^{\overrightarrow{E}} \rightarrow_{\mathrm{v}} F\sigma_0^{\overrightarrow{E}} \dots \sigma_i^{\overrightarrow{E}}[\langle \boldsymbol{E}[E_j\sigma_0^{\overrightarrow{E}} \dots \sigma_{j-1}^{\overrightarrow{E}}[v\sigma_0^{\overrightarrow{E}} \dots \sigma_i^{\overrightarrow{E}}]]\rangle]$$
$$\langle t\rangle\delta_0^{\overrightarrow{E}} \dots \delta_i^{\overrightarrow{E}} \rightarrow_{\mathrm{v}} F\delta_0^{\overrightarrow{E}} \dots \delta_i^{\overrightarrow{E}}[\langle \boldsymbol{E'}[E_j\delta_0^{\overrightarrow{E}} \dots \delta_{j-1}^{\overrightarrow{E}}[v\delta_0^{\overrightarrow{E}} \dots \delta_i^{\overrightarrow{E}}]]\rangle]$$

The resulting terms can be written $\langle s_0^{t',F}\rangle\sigma_0^{\overrightarrow{E}} \dots \sigma_j^{\overrightarrow{E}}$ and $\langle u_0^{t',\overrightarrow{F}}\rangle\delta_0^{\overrightarrow{E}} \dots \delta_j^{\overrightarrow{E}}$, with $t' = E_j[v]$, therefore we obtain terms in $\mathcal{X}_{\mathcal{E}}$.

Let $\langle s_i^{t,\overrightarrow{F}}\rangle\sigma_0^{\overrightarrow{E'}} \dots \sigma_j^{\overrightarrow{E'}} \; \mathcal{X}_{\mathcal{E}} \; \langle u_i^{t,\overrightarrow{F}}\rangle\delta_0^{\overrightarrow{E'}} \dots \delta_j^{\overrightarrow{E'}}$. One can check that the reductions from terms of the form $s_i^{t,\overrightarrow{F}}$, $u_i^{t,\overrightarrow{F}}$ come from respectively $s_0^{t\overrightarrow{F}}$ and $u_0^{t,\overrightarrow{F}}$, and the transitions from these two terms come from t. We have several cases for t. If $t \rightarrow_{\mathrm{v}} t'$, then we still have $\langle s_i^{t',\overrightarrow{F}}\rangle\sigma_0^{\overrightarrow{E'}} \dots \sigma_j^{\overrightarrow{E'}} \; \mathcal{X}_{\mathcal{E}} \; \langle u_i^{t',\overrightarrow{F}}\rangle\delta_0^{\overrightarrow{E'}} \dots \delta_j^{\overrightarrow{E'}}$.

If $t = v$, then $\langle s_0^{v,\overrightarrow{F}}\rangle\sigma_0^{\overrightarrow{E'}} \dots \sigma_j^{\overrightarrow{E'}} = \langle F_0[\langle \boldsymbol{E}[v]\rangle]\rangle\sigma_0^{\overrightarrow{E'}} \dots \sigma_j^{\overrightarrow{E'}}$ and we also have $\langle u_0^{t',\overrightarrow{F}}\rangle\delta_0^{\overrightarrow{E'}} \dots \delta_j^{\overrightarrow{E'}} = \langle F_0[\langle \boldsymbol{E'}[v]\rangle]\rangle\delta_0^{\overrightarrow{E'}} \dots \delta_j^{\overrightarrow{E'}}$. It is easy to check that $\langle \boldsymbol{E}[v]\rangle$ and $\langle \boldsymbol{E'}[v]\rangle$ reduce to the same term $\langle t_0\{\lambda y.\langle E[y]\rangle/k\}\{v/x\}\rangle$, written t'. Then we have $\langle s_0^{v,\overrightarrow{F}}\rangle\sigma_0^{\overrightarrow{E'}} \dots \sigma_j^{\overrightarrow{E'}} \rightarrow_{\mathrm{v}} \langle F_0[t']\rangle\sigma_0^{\overrightarrow{E'}} \dots \sigma_j^{\overrightarrow{E'}}$, and also $\langle u_0^{t',\overrightarrow{F}}\rangle\delta_0^{\overrightarrow{E'}} \dots \delta_j^{\overrightarrow{E'}} \rightarrow_{\mathrm{v}} \langle F_0[t']\rangle\delta_0^{\overrightarrow{E'}} \dots \delta_j^{\overrightarrow{E'}}$; the two resulting terms are in the first set of $\mathcal{X}$. If $i > 0$, one can check that $\langle s_i^{v,\overrightarrow{F}}\rangle\sigma_0^{\overrightarrow{E'}} \dots \sigma_j^{\overrightarrow{E'}} \rightarrow_{\mathrm{v}} \langle s_{i-1}^{\langle F_0[t']\rangle,\overrightarrow{F'}}\rangle\sigma_0^{\overrightarrow{E'}} \dots \sigma_j^{\overrightarrow{E'}}$ and we also have $\langle u_i^{v,\overrightarrow{F}}\rangle\delta_0^{\overrightarrow{E'}} \dots \delta_j^{\overrightarrow{E'}} \rightarrow_{\mathrm{v}} \langle u_{i-1}^{\langle F_0[t']\rangle,\overrightarrow{F'}}\rangle\delta_0^{\overrightarrow{E'}} \dots \delta_j^{\overrightarrow{E'}}$, where $\overrightarrow{F'} = F_1 \dots F_i$ (the first context F_0 is removed from the sequence). We obtain terms that are in the second set of $\mathcal{X}$. In both cases, the resulting terms are in $\mathcal{X}$. The reasoning is the same if $t = k_l$ for some $0 \leq l \leq j$.

If $t = E'_{j+1}[\mathcal{S}k_{j+1}.t']$, then

$$\langle s_0^{t,\overrightarrow{F}}\rangle \sigma_0^{\overrightarrow{E'}} \ldots \sigma_j^{\overrightarrow{E'}} = \langle F_0[\langle \boldsymbol{E}[E'_{j+1}[\mathcal{S}k_{j+1}.t']]\rangle]\rangle \sigma_0^{\overrightarrow{E'}} \ldots \sigma_j^{\overrightarrow{E'}}$$
$$\rightarrow_{\mathrm{v}} \langle F_0[\langle t'\rangle]\rangle \sigma_0^{\overrightarrow{E'}} \ldots \sigma_j^{\overrightarrow{E'}} \sigma_{j+1}^{\overrightarrow{E'},E'_{j+1}}$$

and

$$\langle u_0^{t,\overrightarrow{F}}\rangle \delta_0^{\overrightarrow{E'}} \ldots \delta_j^{\overrightarrow{E'}} = \langle F_0[\langle \boldsymbol{E'}[E'_{j+1}[\mathcal{S}k_{j+1}.t']]\rangle]\rangle \delta_0^{\overrightarrow{E'}} \ldots \delta_j^{\overrightarrow{E'}}$$
$$\rightarrow_{\mathrm{v}} \langle F_0[\langle t'\rangle]\rangle \delta_0^{\overrightarrow{E'}} \ldots \delta_j^{\overrightarrow{E'}} \delta_{j+1}^{\overrightarrow{E'},E'_{j+1}}$$

therefore $\langle s_0^{v,\overrightarrow{F}}\rangle \sigma_0^{\overrightarrow{E'}} \ldots \sigma_j^{\overrightarrow{E'}}$ and $\langle u_0^{v,\overrightarrow{F}}\rangle \delta_0^{\overrightarrow{E'}} \ldots \delta_j^{\overrightarrow{E'}}$ reduce to terms of the form $\langle t''\rangle \sigma_0^{\overrightarrow{E'}} \ldots \sigma_{j+1}^{\overrightarrow{E'}}$ and $\langle t''\rangle \delta_0^{\overrightarrow{E'}} \ldots \delta_{j+1}^{\overrightarrow{E'}}$, that are in $\mathcal{X}_{\mathcal{E}}$. If $i > 0$, then one can check that $\langle s_i^{v,\overrightarrow{F}}\rangle \sigma_0^{\overrightarrow{E'}} \ldots \sigma_j^{\overrightarrow{E'}} \rightarrow_{\mathrm{v}} \langle s_{i-1}^{\langle F_0[t']\rangle,\overrightarrow{F'}}\rangle \sigma_0^{\overrightarrow{E'}} \ldots \sigma_{j+1}^{\overrightarrow{E'}}$ and also $\langle u_i^{v,\overrightarrow{F}}\rangle \delta_0^{\overrightarrow{E'}} \ldots \delta_j^{\overrightarrow{E'}} \rightarrow_{\mathrm{v}} \langle u_{i-1}^{\langle F_0[t']\rangle,\overrightarrow{F'}}\rangle \delta_0^{\overrightarrow{E'}} \ldots \delta_{j+1}^{\overrightarrow{E'}}$, where $\overrightarrow{F'} = F_1 \ldots F_i$, so the resulting terms are in $\mathcal{X}_{\mathcal{E}}$.

If $t = F_{i+1}[k_l\ v]$ (with $1 \leq l \leq j$), then

$$\langle s_0^{t,\overrightarrow{F}}\rangle \sigma_0^{\overrightarrow{E'}} \ldots \sigma_j^{\overrightarrow{E'}} = \langle F_0[\langle \boldsymbol{E}[F_{i+1}[(\lambda y.\langle \boldsymbol{E}[E_l[y]]\rangle)\ v]]\rangle]\rangle \sigma_0^{\overrightarrow{E'}} \ldots \sigma_j^{\overrightarrow{E'}}$$
$$\rightarrow_{\mathrm{v}} \langle F_0[\langle \boldsymbol{E}[F_{i+1}[\langle \boldsymbol{E}[E_l[v]]\rangle]]\rangle]\rangle \sigma_0^{\overrightarrow{E'}} \ldots \sigma_j^{\overrightarrow{E'}} = \langle s_1^{E_l[v],\overrightarrow{F'}}\rangle \sigma_0^{\overrightarrow{E'}} \ldots \sigma_j^{\overrightarrow{E'}}$$

and

$$\langle u_0^{t,\overrightarrow{F}}\rangle \delta_0^{\overrightarrow{E'}} \ldots \delta_j^{\overrightarrow{E'}} = \langle F_0[\langle \boldsymbol{E'}[F_{i+1}[(\lambda y.\langle \boldsymbol{E'}[E_l[y]]\rangle)\ v]]\rangle]\rangle \delta_0^{\overrightarrow{E'}} \ldots \delta_j^{\overrightarrow{E'}}$$
$$\rightarrow_{\mathrm{v}} \langle F_0[\langle \boldsymbol{E'}[F_{i+1}[\langle \boldsymbol{E'}[E_l[v]]\rangle]]\rangle]\rangle \delta_0^{\overrightarrow{E'}} \ldots \delta_j^{\overrightarrow{E'}} = \langle u_1^{E_l[v],\overrightarrow{F'}}\rangle \delta_0^{\overrightarrow{E'}} \ldots \delta_j^{\overrightarrow{E'}}$$

with $\overrightarrow{F'} = F_{i+1}, F_0, \ldots F_i$, so the resulting terms are in $\mathcal{X}_{\mathcal{E}}$. If $i > 0$, then $\langle s_0^{t,\overrightarrow{F}}\rangle \sigma_0^{\overrightarrow{E'}} \ldots \sigma_j^{\overrightarrow{E'}} \rightarrow_{\mathrm{v}} \langle s_{i+1}^{E_l[v],\overrightarrow{F'}}\rangle \sigma_0^{\overrightarrow{E'}} \ldots \sigma_j^{\overrightarrow{E'}}$, and we have also $\langle u_0^{t,\overrightarrow{F}}\rangle \delta_i^{\overrightarrow{E'}} \ldots \delta_j^{\overrightarrow{E'}} \rightarrow_{\mathrm{v}} \langle u_{i+1}^{E_l[v],\overrightarrow{F'}}\rangle \delta_0^{\overrightarrow{E'}} \ldots \delta_j^{\overrightarrow{E'}}$, so the resulting terms are in $\mathcal{X}_{\mathcal{E}}$, as required.

Finally, let $\lambda x.t_0\ \mathcal{E}\ \lambda x.t_1$ and $v_0\ \widehat{\mathcal{E}}\ v_1$. It is easy to check that by definition of $\mathcal{E}$, the two terms $t_0\{v_0/x\}$ and $t_1\{v_1/x\}$ are of the form $t'\sigma_0^{\overrightarrow{E'}} \ldots \sigma_i^{\overrightarrow{E}}$ and $t'\delta_0^{\overrightarrow{E'}} \ldots \delta_i^{\overrightarrow{E}}$.