# RPO, SECOND-ORDER CONTEXTS, AND $\lambda$-CALCULUS

PIETRO DI GIANANTONIO, FURIO HONSELL, AND MARINA LENISA

Dip. di Matematica e Informatica, Università di Udine, via delle Scienze 206, 33100 Udine, Italy
*e-mail address*: {digianantonio,honsell,lenisa}@dimi.uniud.it

Abstract. First, we extend Leifer-Milner RPO theory, by giving general conditions to obtain IPO labelled transition systems (and bisimilarities) with a reduced set of transitions, and possibly finitely branching. Moreover, we study the weak variant of Leifer-Milner theory, by giving general conditions under which the weak bisimilarity is a congruence. Then, we apply such extended RPO technique to the lambda-calculus, endowed with lazy and call by value reduction strategies. We show that, contrary to process calculi, one can deal directly with the lambda-calculus syntax and apply Leifer-Milner technique to a category of contexts, provided that we work in the framework of weak bisimilarities. However, even in the case of the transition system with minimal contexts, the resulting bisimilarity is infinitely branching, due to the fact that, in standard context categories, parametric rules such as the beta-rule can be represented only by infinitely many ground rules. To overcome this problem, we introduce the general notion of second-order context category. We show that, by carrying out the RPO construction in this setting, the lazy (call by value) observational equivalence can be captured as a weak bisimilarity equivalence on a finitely branching transition system. This result is achieved by considering an encoding of lambda-calculus in Combinatory Logic.

## 1. Introduction

Recently, much attention has been devoted to derive *labelled transition systems* and *bisimilarity congruences* from *reactive systems*, in the context of process languages and graph rewriting, [Sew02, LM00, SS03, GM05, BGK06, BKM06, EK06]. In the theory of process algebras, the operational semantics of CCS was originally given via a labelled transition system (lts), while more recent process calculi have been presented via reactive systems plus structural rules. Reactive systems naturally induce behavioral equivalences which are

*2000 ACM Subject Classification:* F.3.2 [Logic and Meaning of Programs]: Semantics of Programming Languages—Algebraic approaches to semantics,Denotational semantics, Operational semantics,Process model; F.4.1 [Mathematical Logic and Formal Languages]: Mathematical Logic — Lambda calculus and related systems.

*Key words and phrases:* Lambda calculus, reactive system, labelled transition system, weak bisimilarity, RPO technique.

Work supported by ART PRIN Project prot. 2005015824 (funded by MIUR).

LOGICAL METHODS
IN COMPUTER SCIENCE

DOI:10.2168/LMCS-???

congruences w.r.t. contexts, while lts's naturally induce bisimilarity equivalences with coinductive characterizations. However, such equivalences are not congruences in general, or else it is an heavy, ad-hoc task to prove that they are congruences.

Generalizing [Sew02], Leifer and Milner [LM00] presented a general categorical method for deriving a transition system from a reactive system, in such a way that the induced bisimilarity is a congruence. The labels in Leifer-Milner's transition system are those contexts which are *minimal* for a given reaction to fire. Minimal contexts are identified via the categorical notion of *relative pushout (RPO)*. Leifer-Milner's central result guaranties that, under a suitable categorical condition, the induced bisimilarity is a *congruence* w.r.t. all contexts.

In the literature, some case studies have been carried out, especially in the setting of process calculi, for testing the expressivity of Leifer-Milner's approach. Some difficulties have arisen in applying the approach directly to such languages, viewed as Lawvere theories, because of structural rules. To overcome this problem, two different approaches have been considered. The first approach consists in using more complex categorical constructions, where structural rules are accounted for explicitly, [Lei01, SS03, SS05]. In the second approach, intermediate encodings have been considered in graph theory, for which the approach of "borrowed contexts" has been developed [EK06], and in Milner's bigraph theory. Here structural rules are avoided, since structurally equivalent terms are equated in the target language.

Moreover, the following further issues have arisen in applying Leifer-Milner's technique.

(i) Leifer-Milner's bisimilarity is still redundant, and many labels have to be eliminated *a posteriori*, by an *ad-hoc* reasoning. Thus general results are called for, in order to reduce the complexity of the bisimilarity *a priori*.

(ii) In some cases it is useful to consider *weak* variants of Leifer-Milner technique. However, for the *weak bisimilarity* we only have a *partial congruence* result, stating that such bisimilarity is a congruence w.r.t. a certain class of contexts. However, in many concrete cases, the weak bisimilarity turn out to be a *full congruence*. Thus it will be useful to study general conditions under which this happens.

(iii) When Leifer-Milner technique is applied in the standard setting of term and context categories (Lawvere theories), the rules in the rewriting system cannot be represented parametrically, but only at a *ground level* through a (infinite) series of possible instantiations. As a consequence, the bisimilarity turns out to be infinitely branching. In [KSS05], a generalization of Leifer-Milner technique for dealing with parametric rules has been introduced. This approach is rather complex and not completely satisfactory. An alternative approach (which is considered in the present paper) consists in studying *second-order* versions of term and context categories, which allow *parametric* representations of rewriting rules, and carrying out Leifer-Milner technique in this setting.

In this paper, we address all the above issues. In particular, in the first part of the paper, we extend Leifer-Milner theory, by providing general results for reducing the complexity of the bisimilarity, and by studying conditions under which the weak bisimilarity is a full congruence. Then, we focus on the prototypical example of reactive system given by the $\lambda$-calculus, endowed with lazy and call by value (cbv) reduction strategies. We show that, in principle, contrary to most of the case studies considered in the literature, one could deal directly with the $\lambda$-calculus syntax and apply Leifer-Milner technique to the category of term contexts induced by the $\lambda$-terms, provided that we work in the setting of *weak* *bisimilarities.* Applying our general results, we get quite economical weak bisimilarities which are congruences and we recover exactly both lazy and cbv contextual equivalences. As a by-product, we also get an alternative proof of the Context Lemma for the lazy case. However, the bisimilarities that we obtain are still infinitely branching. This is mainly due to the fact that, in the category of contexts, the $\beta$-rule cannot be described parametrically, but it needs to be described extensionally using an infinite set of pairs of ground terms. In order to overcome this problem, we consider the combinatory logic and we introduce the general notion of *category of second-order term contexts*, which provide a solution to the third issue above. Our main result amounts to the fact that, by carrying out Leifer-Milner's construction in this setting, the *lazy (cbv) contextual equivalence* can be captured as a *weak bisimilarity equivalence* on a (finitely branching) transition system. Technically, this result is achieved by considering an encoding of the lazy (cbv) $\lambda$-calculus in KS Combinatory Logic (CL), endowed with a lazy (cbv) reduction strategy, and by showing that the lazy (cbv) contextual equivalence on $\lambda$-calculus can be recovered as a lazy (cbv) equivalence on CL. It is necessary to consider such encoding, since the approach of second-order context categories proposed in this paper works for reaction rules which are "local", that is the reaction does not act on the whole term, but only locally. But the substitution operation on $\lambda$-calculus is not local.

Finally, the correspondence results obtained in this paper about the observational equivalences on $\lambda$-calculus and CL are interesting *per se* and, although natural and ultimately elementary, had not appeared previously in the literature.

*Summary.* In Section 2, we summarize the theory of reactive systems of [LM00]. In Section 3, we extend such theory with new general results about weak bisimilarity, and about the "pruning" of Leifer-Milner lts and the induced bisimilarity. In Section 4, we present the $\lambda$-calculus together with lazy and cbv reduction strategies and observational equivalences, and we discuss the RPO approach applied to the $\lambda$-calculus endowed with a structure of context category. In Section 5, we focus on Combinatory Logic (CL), we show how to recover on CL the lazy and cbv strategies and observational equivalences, and we discuss the RPO approach applied to CL, viewed as a context category. In Section 6, we introduce the notion of second-order context category, and we apply the RPO approach to CL viewed as a second-order rewriting system, thus obtaining characterizations of lazy and cbv observational equivalences as weak bisimilarities on finitely branching lts's. Final remarks and directions for future work appear in Section 7.

The present paper extends [DHL08]. The main new contribution of the present paper is the extension of Leifer-Milner theory, which appears in Section 3. This allows to deal with the $\lambda$-calculus in the subsequent sections in a smoother way, to get stronger results about the lts and the induced bisimilarity, both for the lazy and for the cbv case, and also to provide an alternative proof of the Context Lemma in the lazy case.

## 2. The Theory of Reactive Systems

In this section, we summarize the theory of reactive systems proposed in [LM00] to derive lts's and bisimulation congruences from a given reduction semantics. Moreover, we discuss weak variants of Leifer-Milner's bisimilarity equivalence.

The theory of [LM00] is based on a categorical formulation of the notion of *reactive system*, whereby *contexts* are modeled as arrows of a category, *terms* are arrows having as domain 0 (a special object which denotes no holes), and reaction rules are pairs of terms.

**Definition 2.1** (Reactive System)**.** A *reactive system* $\mathbf{C}$ consists of:

- a category $\mathcal{C}$;
- a distinguished object $0 \in |\mathcal{C}|$;
- a composition-reflecting subcategory $\mathcal{D}$ of *reactive contexts*;
- a set of pairs $\mathbf{R} \subseteq \bigcup_{I\in|\mathcal{C}|} \mathcal{C}[0,I] \times \mathcal{C}[0,I]$ of *reaction rules*.

The reactive contexts are those in which a reaction can occur. By composition-reflecting we mean that $dd' \in \mathcal{D}$ implies $d, d' \in \mathcal{D}$.

Reactive systems on term languages can be viewed as a special case of reactive systems in the sense of Leifer-Milner by instantiating $\mathcal{C}$ as a suitable category of term and contexts, also called the (free) Lawvere category, [LM00]. In this view, we often call terms the arrows with domains 0, and contexts the other arrows.

From the set of reaction rules one generates the reaction relation by closing them under all reactive contexts:

**Definition 2.2** (Reaction Relation)**.** Given a reaction system with reactive contexts $\mathcal{D}$ and reaction rules $\mathbf{R}$, the *reaction relation* $\rightarrow$ is defined by:

$$t \rightarrow u \quad \text{iff} \quad t = dl,\ u = dr \text{ for some } d \in \mathcal{D} \text{ and } \langle l, r\rangle \in \mathbf{R}\ .$$

The behavior of a reactive system is expressed as an unlabelled transition system. On the other hand, many useful behavioral equivalences are only defined for lts's. The passage from reactive systems to lts's is obtained as follows.

**Definition 2.3** (Context Labelled Transition System)**.** Given a reactive system $\mathbf{C}$, the associated context lts is defined as follows:

- states: arrows $t : 0 \rightarrow I$ in $\mathcal{C}$, for any $I$;
- transitions: $t \xrightarrow{c}_C u$ iff $c \in \mathcal{C}$ and $ct \rightarrow u$.

In the case of a reactive system defined on a category of contexts, a state is a term $t$, and an associated label is a context $c$ such that $ct$ reduces. In the following, we will consider also lts's obtained by reducing the set of transitions of the context lts. In the sequel, we will use the word lts to refer to any such lts obtained from a context lts.

Any lts induces a bisimilarity relation as follows:

**Definition 2.4** (Bisimilarity)**.** Let $\xrightarrow{c}$ be a lts.

(i) A symmetric relation $\mathcal{R} \subseteq \bigcup_{I\in C} \mathcal{C}(0,I) \times \mathcal{C}(0,I)$ on the states of the lts is a *bisimulation* if: $\langle a, b\rangle \in \mathcal{R} \ \wedge\ a \xrightarrow{f} a' \implies \exists b'.\ b \xrightarrow{f} b' \ \wedge\ \langle a', b'\rangle \in \mathcal{R}$ .

(ii) We call *bisimilarity* the largest bisimulation.

(iii) The bisimilarity on the context lts is called *context bisimilarity* $\sim_C$.

It is easy to check that the context bisimilarity is a *congruence* w.r.t. all contexts, i.e. if $a \sim_C b$, then for any context $c$, $ca \sim_C cb$. However, intuitively only those contexts which contain the *minimal* amount of information for a reaction to fire are relevant, while the others are redundant. Moreover, often context bisimilarity gives an equivalence which is too coarse, as we will see also in this paper. Thus, in [LM00], the authors proposed a categorical criterion for identifying the "smallest context allowing a reaction". They defined *relative*

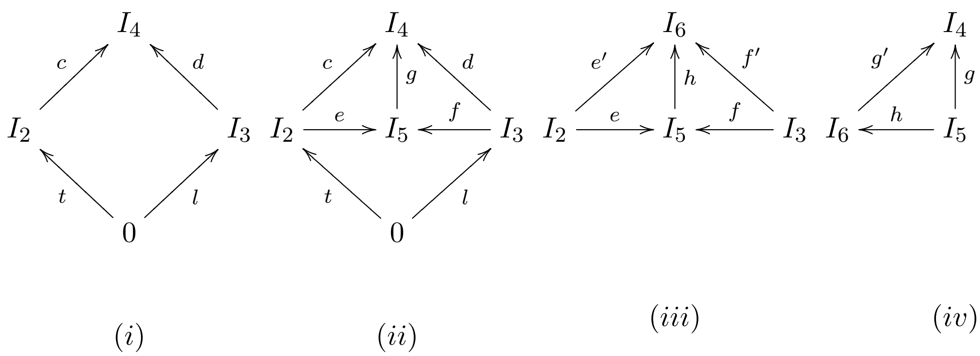



Figure 1: Redex Square and Relative Pushout.

*pushouts* (RPOs), of which *idem relative pushouts* (IPOs) are a special case. One can define a lts using IPOs. Leifer-Milner's central result consists in showing that, under a suitable categorical condition, such lts is well-behaved, in the sense that the induced bisimilarity is a congruence.

**Definition 2.5** (RPO/IPO)**.**

(i) Let $\mathcal{C}$ be a category and let us consider the commutative diagram in Fig. 1(i). Any tuple $\langle I_5, e, f, g\rangle$ which makes diagram in Fig. 1(ii) commute is called a *candidate* for (i). A *relative pushout (RPO)* is the smallest such candidate, i.e. it satisfies the universal property that given any other candidate $\langle I_6, e', f', g'\rangle$, there exists a *unique* mediating morphism $h : I_5 \to I_6$ such that both diagrams in Fig. 1(iii) and Fig. 1(iv) commute.

(ii) A commuting square such as diagram in Fig 1(i) is an *idem pushout (IPO)* if $\langle I_4, c, d, id_{I_4}\rangle$ is its RPO.

**Definition 2.6** (IPO Transition System)**.**

- States: arrows $t : 0 \to I$ in $\mathcal{C}$, for any $I$;
- Transitions: $t \xrightarrow{c}_I dr$ iff $d \in \mathcal{D}$, $ct = dl$, $\langle l, r\rangle \in \mathbf{R}$ and the diagram in Fig. 1(i) is an IPO.

Let $\sim_I$ denote the bisimilarity induced by the IPO lts.

**Definition 2.7** (Redex Square)**.** Let $\mathbf{C}$ be a reactive system and $t : 0 \to I_2$ an arrow in $\mathcal{C}$. A *redex square* (see Fig. 1(i)) consists of a left-hand side $l : 0 \to I_3$ of a reaction rule $\langle l : 0 \to I_3, r : 0 \to I_3\rangle \in \mathbf{R}$, a context $c : I_2 \to I_4$ and a reactive context $d : I_3 \to I_4$ such that $ct = dl$.
A reactive system $\mathbf{C}$ is said to *have redex RPOs* if every redex square has an RPO.

The following is a fundamental lemma stating a property of IPO squares.

**Lemma 2.8** (IPO pasting, [LM00])**.** *Suppose that the square in Fig. 2(i) has an RPO and that both squares in Fig. 2(ii) commute.*

(i) *If the two squares of Fig. 2(ii) are IPOs so is the outer rectangle.*

(ii) *It the outer rectangle and the left square of Fig. 2(ii) are IPOs so is the right square.*

From the above lemma Leifer and Milner derived their central result:

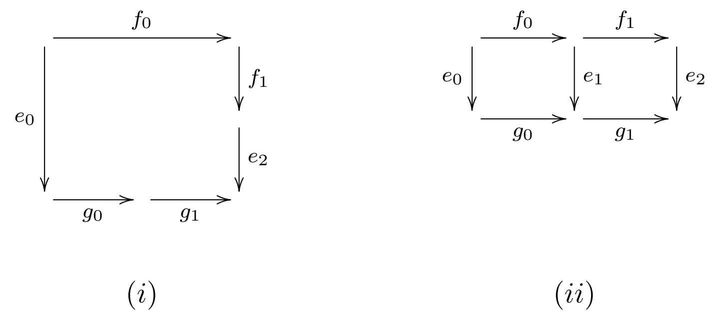



Figure 2: IPO pasting.

**Theorem 2.9** ([LM00])**.** *Let* $\mathbf{C}$ *be a reactive system having redex RPOs. Then the IPO bisimilarity* $\sim_I$ *is a congruence w.r.t. all contexts, i.e. if* $a\sim_I b$ *then for all* $c$ *of the appropriate type,* $ca\sim_I cb$*.*

2.1. **Weak Bisimilarity.** For dealing with the $\lambda$-calculus, it will be useful to consider the weak versions of the context and IPO lts's defined above, together with the corresponding notions of *weak bisimilarities*.

One can proceed in general, by defining a weak lts from a given lts:

**Definition 2.10** (Weak lts and Bisimilarity)**.** Let $\stackrel{\alpha}{\longrightarrow}$ be a lts, and let $\tau$ be a label (identifying an unobservable action).

(i) We define the *weak lts* $\stackrel{\alpha}{\Longrightarrow}$ by

$$t \stackrel{\alpha}{\Longrightarrow} u \;\text{ iff }\; \begin{cases} t \stackrel{\tau}{\longrightarrow}^* u & \text{if } \alpha = \tau \\ t \stackrel{\tau}{\longrightarrow}^* t' \stackrel{\alpha}{\longrightarrow} u' \stackrel{\tau}{\longrightarrow}^* u & \text{otherwise} \end{cases},$$

where $\stackrel{\tau}{\longrightarrow}^*$ denotes the reflexive and transitive closure of $\stackrel{\tau}{\longrightarrow}$.

(ii) Let us call *weak bisimilarity* the bisimilarity induced by the weak lts.

The above definition differs from the one proposed in [LM00]. We cannot use the latter, since it discriminates $\lambda$-terms which are equivalent in the usual semantics.

The following easy lemma gives a useful characterization of the weak bisimilarity, whereby any $\stackrel{\alpha}{\longrightarrow}$-transition is mimicked by a $\stackrel{\alpha}{\Longrightarrow}$-transition:

**Lemma 2.11.** *Let* $\stackrel{\alpha}{\longrightarrow}$ *be a lts and let* $\stackrel{\alpha}{\Longrightarrow}$ *be the corresponding weak lts. The induced weak bisimilarity is the greatest symmetric relation* $\mathcal{R}$ *s.t.:*

$$\langle a, b\rangle \in \mathcal{R} \;\wedge\; a \stackrel{f}{\longrightarrow} a' \;\Longrightarrow\; \exists b'.\, b \stackrel{f}{\Longrightarrow} b' \;\wedge\; \langle a', b'\rangle \in \mathcal{R}\ .$$

The following lemma provides a coinduction "up-to" principle, which will be useful in the sequel:

**Lemma 2.12.** *Let* $\stackrel{\alpha}{\longrightarrow}$ *be a lts and let* $\stackrel{\alpha}{\Longrightarrow}$ *be the corresponding weak lts. The induced weak bisimilarity is the greatest symmetric relation* $\mathcal{R}$ *s.t.:*

$$\langle a, b\rangle \in \mathcal{R} \;\wedge\; a \stackrel{f}{\Longrightarrow}' a' \;\Longrightarrow\; \exists b'.\, b \stackrel{f}{\Longrightarrow} b' \;\wedge\; \langle a', b'\rangle \in \mathcal{R}^*\ ,$$

*where* $\stackrel{f}{\Longrightarrow}'$ *denotes* $\stackrel{\tau}{\longrightarrow}^* \circ \stackrel{f}{\longrightarrow}$ *(*$f$ *is possibly* $\tau$*), and* $\mathcal{R}^*$ *denotes the reflexive and transitive closure of* $\mathcal{R}$*.*

*Proof.* Let us call "bisimulation up-to" a relation $\mathcal{R}$ as in the statement of the lemma. In order to prove the thesis, it is sufficient to prove that, if $\mathcal{R}$ is a bisimulation up-to, then $\mathcal{R}^*$ is a bisimulation. Let $\mathcal{R}$ be a bisimulation up-to. First, one can easily check that $(a\mathcal{R}^*b \ \wedge \ a \Longrightarrow a') \ \Longrightarrow \ \exists b'. \ (b \Longrightarrow b' \ \wedge \ a'\mathcal{R}^*b')$ (by induction on the length of the chain $a \ \mathcal{R} \ldots \mathcal{R} \ b$). Now, let $a = a_0 \ \mathcal{R} \ a_1 \ldots a_{n-1} \ \mathcal{R} \ a_n = b$ and $a \stackrel{f}{\Longrightarrow} a'$. We prove that $\exists b'. \ (b \stackrel{f}{\Longrightarrow} b' \ \wedge \ a'\mathcal{R}^*b')$, by induction on $n \geq 0$. If $n = 0$, the thesis is immediate. If $n > 0$ and $a \stackrel{f}{\Longrightarrow}{}' \ a'' \Longrightarrow a'$, then, since $\mathcal{R}$ is a bisimulation up-to, $a_1 \stackrel{f}{\Longrightarrow} a_1'' \ \wedge \ a''\mathcal{R}^*a_1''$, and, by what we have proved before, $\exists a_1'. \ (a_1'' \Longrightarrow a_1' \ \wedge \ a'\mathcal{R}^*a_1')$. Finally, by induction hypothesis, $\exists b'. \ (b \stackrel{f}{\Longrightarrow} b' \ \wedge \ a_1'\mathcal{R}^*b')$. Hence $a'\mathcal{R}^*b'$. □

For dealing with the $\lambda$-calculus, we will consider a notion of *weak IPO bisimilarity*, where the identity context is unobservable. Such notions of weak IPO bisimilarities are not congruences w.r.t. all contexts, in general, however, as observed in [LM00] (end of Section 5), they are congruences at least w.r.t. reactive contexts:

**Theorem 2.13.** *Let* **C** *be a reactive system having redex RPOs. Then the weak IPO bisimilarity* $\approx_I$*, where the identity context is unobservable, is a congruence w.r.t. reactive contexts.*

## 3. Extending the Theory of Reactive Systems

In this section, we present some original results concerning the lts obtained by the RPO construction. These results concern two issues:

**Weak-bisimilarity:** Since in the $\lambda$-calculus the weak bisimilarity is the equivalence to be used, we present some general conditions assuring that the weak bisimilarity, on the lts obtained by an IPO construction, is a congruence w.r.t. all contexts.

**Pruning the lts tree:** In order to obtain a feasible lts, i.e. a lts with a reduced set of transitions, possibly finitely branching, it is often necessary to prune the lts obtained by an IPO construction. We present some general conditions allowing to prune IPO lts', without modifying the induced (weak)-bisimilarity.

We present our results in two different versions, the first one is quite simple, but it does not apply to our particular case, so we present a second version that is more involved but suits our needs. We choose to present the simple first version of our results as an introduction to the second one, and also because it can have applications in modeling languages different from the $\lambda$-calculus.

Some preliminary definitions are necessary. F

**Definition 3.1.** Let $\mathbb{C}$ be a reactive system having redex RPOs:

- Given a set of IPO labels $L$, the *$L$-restricted IPO lts* is the lts obtained by removing from the IPO lts all transitions not labeled by elements in $L$. We denote by $\approx_L$ the weak bisimilarity induced by the $L$-restricted IPO lts.
- We denote by $R$ the set of IPO labels that are reactive contexts. We denote by $\approx_R$ the weak bisimilarity induced by the $R$-restricted IPO lts.
- In a reactive system, we say that the family of IPO transitions with label $f : I_0 \to I_1$ is *definable by contexts* if there exists a list of contexts $e_1, \ldots, e_h : I_0 \to I_1$ such that, for all $t : 0 \to I_0, t' : 0 \to I_1$, we have that $t \stackrel{f}{\longrightarrow}_I t'$ iff $\exists i. \ t' = e_i t$.

Intuitively, a family of IPO transitions with label $f : I_0 \to I_1$ is definable by contexts if $f$ is an IPO for any arrow $t : 0 \to I_0$ and the IPO transitions on $f$ can be described by contexts, that is, they do not modify the internal structure of the term $t$.

**Proposition 3.2.** *Let* **C** *be a reactive system having redex RPOs. If any IPO context is either reactive or definable by contexts (or both), then the weak IPO bisimilarity $\approx_I$ (with the identity IPO context unobservable) is a congruence. Moreover $\approx_I$ coincides with $\approx_R$.*

*Proof.* Consider the relation $S = \{\langle ct, cu\rangle \mid t \approx_R u,\ c \text{ context}\}$. It is immediate that $\approx_I \subseteq \approx_R$, and from this, $\approx_I \subseteq \{\langle ct, cu\rangle \mid t \approx_I u,\ c \text{ context}\} \subseteq S$. If we prove also the inclusion $S \subseteq \approx_I$, then all relations are equal and $\approx_I$ coincides with its contextual closure, i.e. it is congruence. By Lemma 2.11, in order to prove $S \subseteq \approx_I$ it is sufficient to show that, for any $\langle ct, cu\rangle \in S$, if $ct \xrightarrow{f}_I t'$ then there exists $u'$ s.t. $cu \stackrel{f}{\Longrightarrow}_I u'$ with $t' S u'$.

Consider the following diagram:

$$\begin{array}{ccccc} 0 & \xrightarrow{t} & I_0 & \xrightarrow{c} & I_2 \\ {\scriptstyle l}\downarrow & & {\scriptstyle f'}\downarrow & & \downarrow{\scriptstyle f} \\ I_3 & \xrightarrow[d]{} & I_1 & \xrightarrow[d']{} & I_4 \end{array}$$

where the outermost rectangle is the IPO inducing the transition $ct \xrightarrow{f}_I t'$, namely $t' = d'dr$ with $\langle l, r\rangle$ a reaction rule, while the left square is obtained from an IPO construction starting from $l$ and $t$. By Lemma 2.8, the IPO pasting property, we have that also the right-hand square of the diagram is an IPO.

There are two cases to consider:

(i) If the context $f'$ is definable by contexts, since $t \xrightarrow{f'}_I dr$, there exists a context $e$ such that $dr = et$ and $t' = d'et$, it follows that $u \xrightarrow{f'}_I eu$. That is there exist a reaction rule $\langle l_1, r_1\rangle$ and a reactive context $d_1$ s.t. $eu = d_1 r_1$, and the left-hand square of the following diagram is a IPO.

$$\begin{array}{ccccc} 0 & \xrightarrow{u} & I_0 & \xrightarrow{c} & I_2 \\ {\scriptstyle l_1}\downarrow & & {\scriptstyle f'}\downarrow & & \downarrow{\scriptstyle f} \\ I_3 & \xrightarrow[d_1]{} & I_1 & \xrightarrow[d']{} & I_4 \end{array}$$

Since the right-hand square is IPO, by the IPO pasting property, Lemma 2.8, also the outermost rectangle is an IPO. It follows that $cu \xrightarrow{f}_I d'd_1 r_1 = d'eu$, from which the thesis.

(ii) If the context $f'$ is reactive, then it so also the context $d'f'$ (composition of reactive contexts) and the context $c$ (reactive contexts are composition-reflecting). Moreover, by the definition of bisimilarity, there exists $u_0$ such that $u \stackrel{f'}{\Longrightarrow}_I u_0$ (i.e. $u \xrightarrow{Id}_I^{\ *} u_1 \xrightarrow{f'}_I u_2 \xrightarrow{Id}_I^{\ *} u_0$) with $u_0 \approx_R dr$. Since $c$ is reactive and squares of the form

$$\begin{array}{ccc} I_0 & \xrightarrow{c} & I_2 \\ {\scriptstyle Id}\downarrow & & \downarrow{\scriptstyle Id} \\ I_1 & \xrightarrow[c]{} & I_3 \end{array}$$

are IPOs, by composition of IPO squares (and by induction) it is easy to prove that $cu \xrightarrow{Id}_I^* cu_1 \xrightarrow{f}_I d'u_2 \xrightarrow{Id}_I^* d'u_0$, from which the thesis.

□

In dealing with the $\lambda$-calculus, we need a more complex version of the above proposition, in which we consider both the category of unary linear term contexts and a category of "multi-holed" linear term contexts. The category of unary contexts is the most suitable for the IPO construction, while in extending the notion of "definable by contexts" it is convenient to consider multi-holed contexts.

The following definition formalizes the relation existing between the two categories of contexts.

**Definition 3.3.** A category $\mathcal{D}$ is a *list extension* of a category $\mathcal{C}$ if the following hold:

- $\mathcal{C}$ contains a distinguished object 0.
- The objects of $\mathcal{D}$ are finite lists of objects of $C$ different from 0.
- By identifying 0 with the empty list $\langle\ \rangle$, and any other object $I$ in $\mathcal{C}$ with the singleton list $\langle I \rangle$, $\mathcal{C}$ is a *full subcategory* of $\mathcal{D}$.
- There exists a *concatenation* functor $\otimes$ from $\mathcal{D} \times \mathcal{D}$ to $\mathcal{D}$ acting as concatenation on objects $\langle I_0, \ldots, I_n \rangle \otimes \langle J_0, \ldots, J_m \rangle = \langle I_0, \ldots, I_n, J_0, \ldots, J_m \rangle$ and being associative on arrows.

In the spirit of the previous remark we will call unary (single-holed) contexts the arrows in $\mathcal{C}$ (with domain different from 0), and multi-holed contexts the arrows in $\mathcal{D}$.

Two other definitions are necessary.

**Definition 3.4.** Given a reactive system $\mathbf{C}$ on a category $\mathcal{C}$, and a category $\mathcal{D}$, list extension of $\mathcal{C}$:

(i) we define a multi-holed context $g : \langle I_0, \ldots, I_n \rangle \to I$ *IPO uniform* if for any IPO context $f$ (in $\mathbf{C}$), there exists a list of multi-holed contexts $g_1 : \langle I_{1,0}, \ldots, I_{1,n_1} \rangle \to I, \ldots, g_h : \langle I_{h,0}, \ldots, I_{h,n_h} \rangle \to I$, and a list of functions $l_1 : \{0, \ldots, n_1\} \to \{0, \ldots, n\}$, $\ldots, l_h : \{0, \ldots, n_h\} \to \{0, \ldots, n\}$ such that: for any n-tuple of $\mathcal{C}$ terms $t_0 : 0 \to I_0, \ldots, t_n : 0 \to I_n$, we have: $g(t_0 \otimes \ldots \otimes t_n) \xrightarrow{f}_I t'$ iff there exists $i$ such that $t' = g_i(t_{l_i(0)} \otimes \ldots \otimes t_{l_i(n_i)})$

(ii) a context $g : \langle I_0, \ldots, I_n \rangle \to I$ has a *reactive index* $i$ if for any list of $n$ terms $t_0 : 0 \to I_0, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n : 0 \to I_n$, the context $f(t_0 \otimes \ldots \otimes t_{i-1} \otimes id_{I_i} \otimes t_{i+1} \otimes \ldots \otimes t_n) : I_i \to I$, seen as a context in $\mathcal{C}$, is reactive.

Intuitively, a context $g$ is IPO uniform if the behaviour wrt the IPO reaction of the term $g(t_{l_i(0)} \otimes \ldots \otimes t_{l_i(n_i)})$ does not depend on the terms $t_{l_i(0)}, \ldots, t_{l_i(n_i)}$.

**Proposition 3.5.** *Let $\mathbf{C}$ be a reactive system having redex RPOs.*

(i) *The weak IPO bisimilarity $\approx_I$ (with the identity IPO context unobservable) is a congruence if there exists a category $\mathcal{D}$, list extension of $\mathcal{C}$ such that any (multi-holed) context $g : \langle I_0, \ldots, I_n \rangle \to I$ is either IPO uniform or it has a reactive index (or both).*

(ii) *Moreover, if the reaction relation is deterministic, i.e. any term can react in at most one possible way, then the relation $\approx_I$ coincides with $\approx_R$.*

*Proof.* Here we present only the proof of point (2). The proof of point (1) is almost identical and can be derived, from the present proof, by substituting the relation $\approx_R$ with $\approx_I$, and by simplifying some steps.

By repeating the same arguments used at the beginning of the proof of Proposition 3.2, it is sufficient to prove that the relation $S = \{\langle\ g(t_0 \otimes \ldots \otimes t_n), g(u_0 \otimes \ldots \otimes u_n)\rangle\ \ |\ g : \langle I_0, \ldots, I_n\rangle \to I, \forall i.t_i \approx_R u_i\}$ is contained in the weak bisimilarity. By Lemma 2.12, it is sufficient to show that for any $\langle g(t_0 \otimes \ldots \otimes t_n), g(u_0 \otimes \ldots \otimes u_n)\rangle \in S$ and $f$ IPO-transition, if $g(t_0 \otimes \ldots \otimes t_n) \stackrel{f}{\Longrightarrow}_I t$, with $\stackrel{f}{\longrightarrow}_I$ the last step of the chain of reactions, then there exists $u$ s.t. $g(u_0 \otimes \ldots \otimes u_n) \stackrel{f}{\Longrightarrow}_I u$ with $tS^*u$. The proof is by double induction on the number of steps of the transition $g(t_0 \otimes \ldots \otimes t_n) \stackrel{f}{\Longrightarrow}_I t$, and on the number $n$ of holes in the list context $g$.

The basic case is when $g(t_0 \otimes \ldots \otimes t_n) \stackrel{f}{\Longrightarrow}_I t$ in 0 steps ($f = id$), in this case there is nothing to prove.

Let suppose $g(t_0 \otimes \ldots \otimes t_n) \stackrel{f'}{\longrightarrow}_I t' \stackrel{f''}{\Longrightarrow}_I t$, in this case $(f' = id \wedge f'' = f)$ or $(f' = f \wedge f'' = id \wedge t' = t)$,

There are two cases to consider:

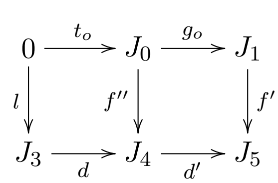
(i) The context $g$ is IPO-uniform: in this case there exists a context $e : \langle I'_0, \ldots, I'_{n'}\rangle \to J_1$ and a function $l : \{0, \ldots n'\} \to \{0, \ldots, n\}$ such that $t' = e(t_{l(0)} \otimes \ldots \otimes t_{l(n')})$ and $g(u_0 \otimes \ldots \otimes u_n) \stackrel{f'}{\longrightarrow}_I e(u_{l(0)} \otimes \ldots \otimes u_{l(n')})$. By application of the inductive hypothesis, on a smaller number of transitions steps, there exists $u$ s.t. $e(u_{l(0)} \otimes \ldots \otimes u_{l(n')}) \stackrel{f''}{\Longrightarrow}_I u$ with $tS^*u$, and from which the thesis.

(ii) The context $g$ has a reactive index $i$, for the sake of simplicity, assume $i = 0$. Consider the arrow $g' = g(t_0 \otimes id_{I_1} \otimes \ldots \otimes id_{I_n}) : \langle I_1, \ldots, I_n\rangle \to I$. Since $g'(t_1 \otimes \ldots \otimes t_n) = g(t_0 \otimes \ldots \otimes t_n) \stackrel{f}{\Longrightarrow}_I t$, by inductive hypothesis, on the number of holes in the multi-holed contexts, there exists $u$ such that $g'(u_1 \otimes \ldots \otimes u_n) = g(t_0 \otimes u_1 \otimes \ldots \otimes u_n) \stackrel{f}{\Longrightarrow}_I u$, with $tS^*u$.

Now consider the context $g'' = g(Id_{I_0} \otimes u_1 \otimes \ldots \otimes u_n) : I_o \to I$. The context $g''$ is reactive and $g''(t_0) \stackrel{f}{\Longrightarrow}_I u$. To obtain the thesis it remains to prove that there exists $u'$ s.t. $g''(u_0) = g(u_0 \otimes \ldots \otimes u_n) \stackrel{f}{\Longrightarrow}_I u'$, with $uS^*u'$.

More generally we prove that for any reactive context $g_o : J_0 \to J_1$, any IPO context $f : J_1 \to J_2$, and any pair of terms $t_o, u_o$, if $t_o \approx_R u_o$ and $g_o(t_o) \stackrel{f}{\Longrightarrow}_I t'_o$ then there exists $u'_o$ s.t. $g_o(u_o) \stackrel{f}{\Longrightarrow}_I u'_o$ and $t'_o S^* u'_o$. The proof is by induction on the number of steps in the transition $g''(t_0) \stackrel{f}{\Longrightarrow}_I t'_o$. The basic case is when the reaction is of zero steps; in this case there is nothing to prove.

For the inductive case consider the following diagram of IPO squares defining the first reaction in the chain



We need to consider two cases. The first one is where $f'$ is a reactive context ($f' \in \{f, Id\}$). Since reactive contexts are composition-reflecting, then also the IPO context $f''$ is reactive. By the definition of bisimilarity, $u_o \stackrel{f''}{\Longrightarrow}_I u_i$ with $u_i \approx_R dr$. By reactivity of $g_o$, using suitable IPO pasting diagrams, we can prove $g_o(u_o) \stackrel{f'}{\Longrightarrow}_I d' u_i$. Now by applying the inductive hypothesis to the reduction $d'(dr) \stackrel{f}{\Longrightarrow}_I t'_o$, we obtain the thesis.

The second case where $f'$, is a non reactive context ($f' = f$). Since reactive contexts are compositional reflecting, then also the IPO context $f''$ is non reactive and therefore, by hypothesis, IPO uniform. Notice that the context $Id$ is an IPO context for the term $f''(t_o)$, by the IPO uniformity of $f''$, $Id$ is an IPO context also for $f''(u_o)$ and there exist a list context $g'\langle J_1, \dots J_1\rangle \to J$ s.t. $t_o \stackrel{f''}{\longrightarrow}_I g'(t_o, \dots, t_o)$ and also $u_o \stackrel{f''}{\longrightarrow}_I g'(u_o, \dots, u_o)$. Notice that, if the reduction relation is deterministic, two terms that reduce one to the other by $\tau$ transitions are weakly bisimilar, it follows that $g_o(t_o) \stackrel{f}{\longrightarrow}_I d'g'(t_0, \dots, t_0) \approx_R t'_o$ and, by IPO pasting, $g_o(u_o) \stackrel{f}{\longrightarrow}_I d'g'(u_o, \dots, u_o)$, from which we derive the thesis.

□

**Remark 3.6.** Propositions 3.2 and 3.5 above, about congruence of the weak IPO bisimilarity, are more related than what they look at first glance. From one side, by exploiting the fact that the composition of a non-reactive context with any context gives a non-reactive context, one can show that if the non-reactive IPOs are definable by context then any non reactive context is IPO-uniform. Note that the condition of "definability by context" is in general simpler to verify than the one of "IPO-uniformity", and so we prefer to present the given formulation of proposition. On the other side, it would be possible to extend the notion of "definability by context" to the case of list extension categories, however to this aim it would be necessary to present a series of new definitions, necessary to lift the IPO construction to the list extension categories. For the sake of simplicity, we prefer to avoid the introduction of these further notions.

## 4. The Lambda Calculus

First, we recall the $\lambda$-calculus syntax together with *lazy* and *cbv* reduction strategies and observational equivalences. Then, we show how to apply the RPO technique to $\lambda$-calculus, viewed as a context category, and we discuss some problematic issues.

### 4.1. **Syntax, Reduction Strategies, Observational Equivalences.**

**Definition 4.1** (Syntax)**.** The set of *$\lambda$-terms* $\Lambda$ is defined by

$$(\Lambda \ni)\ M ::=\ x \mid MM \mid \lambda x.M\ ,$$

where $x \in Var$ is an infinite set of variables.
Let $FV(M)$ denote the set of free variables in $M$, and let us denote by $\Lambda^0$ the set of *closed* $\lambda$-terms.

As usual, $\lambda$-terms are taken up-to $\alpha$-conversion, and application associates to the left. We consider the standard notions of $\beta$-rule and $\beta_V$-rule:

**Definition 4.2.**

(i) $\beta$-rule: $(\lambda x.M)N \to_\beta M[N/x]$;

(ii) $\beta_V$-rule: $(\lambda x.M)N \to_{\beta_V} M[N/x]$, if $N$ is a variable or a $\lambda$-abstraction.

As usual, we denote by $=_\beta$ and $=_{\beta_V}$ the corresponding conversions.

A *reduction strategy* on the $\lambda$-calculus determines, for each term which is not a value, a suitable $\beta$-redex appearing in it to be contracted. The lazy and cbv reduction strategies are defined on closed $\lambda$-terms as follows:

**Definition 4.3** (Reduction Strategies)**.**

(i) The *lazy strategy* $\to_l \subseteq \Lambda^0 \times \Lambda^0$ reduces the leftmost $\beta$-redex, not appearing within a $\lambda$-abstraction. Formally, $\to_l$ is defined by the rules:

$$\frac{}{(\lambda x.M)N \to_l M[N/x]} \qquad \frac{N \to_l N'}{NP \to_l N'P}$$

(ii) The *call by value strategy* $\to_v \subseteq \Lambda^0 \times \Lambda^0$ reduces the leftmost $\beta_V$-redex, not appearing within a $\lambda$-abstraction. Formally, $\to_v$ is defined by the following rules:

$$\frac{}{(\lambda x.M)V \to_v M[V/x]} \qquad \frac{N \to_v N'}{NP \to_v N'P} \qquad \frac{N \to_v N'}{(\lambda x.M)N \to_v (\lambda x.M)N'}$$

where $V$ is a closed value, i.e. a $\lambda$-abstraction.

We denote by $\to_\sigma^*$ the reflexive and transitive closure of a strategy $\to_\sigma$, for $\sigma \in \{l, v\}$, by $Val_\sigma$ the set of values, i.e. the set of terms on which the reduction strategy halts (which coincides with the set of $\lambda$-abstractions in both cases), and by $M \Downarrow_\sigma$ the fact that there exists $V \in Val_\sigma$ such that $M \to_\sigma^* V$.

As we will see in Section 4.2 below, each strategy defines a (deterministic) reactive system on $\lambda$-terms in the sense of Definition 2.1. To this aim, it is useful to notice that the above reduction strategies can be alternatively determined by specifying suitable sets of *reactive contexts* (see Remark 4.5 below), which are subsets of the following *unary contexts*, i.e. contexts with a single hole:

**Definition 4.4** (Unary Contexts)**.** Let $P \in \Lambda$.

$$C[\,] \ ::= \ [\,] \mid PC[\,] \mid C[\,]P \mid \lambda x.C[\,] \ .$$

**Remark 4.5.**

(i) The lazy strategy $\to_l$ is the closure of the $\beta$-rule under the reactive contexts, corresponding to the closed applicative contexts: $D[\,] \ ::= \ [\,] \mid D[\,]P$ , where $P \in \Lambda^0$.

(ii) The cbv strategy $\to_v$ is the closure of the $\beta_V$-rule under the following closed reactive contexts: $D[\,] \ ::= \ [\,] \mid D[\,]P \mid (\lambda x.M)D[\,]$ , where $P, \lambda x.M \in \Lambda^0$.

Each strategy induces an *observational (contextual) equivalence* à la Morris on closed terms, when we consider programs as *black boxes* and only observe their "halting properties".

**Definition 4.6** ($\sigma$-observational Equivalence)**.** Let $\to_\sigma$ be a reduction strategy and let $M, N \in \Lambda^0$. The *observational equivalence* $\approx_\sigma$ is defined by

$$M \approx_\sigma N \ \text{ iff for any closed unary context } C[\,].\ C[M] \Downarrow_\sigma \Leftrightarrow C[N] \Downarrow_\sigma \ .$$

The definition of $\approx_\sigma$ can be extended to open terms by considering closing (by-value) substitutions, i.e. for $M, N \in \Lambda$ s.t. $FV(M, N) \subseteq \{x_1, \ldots, x_n\}$, we define:

$$M \widehat{\approx}_\sigma N \text{ iff for all closing (by-value) substitutions } \vec{P},\ M[\vec{P}/\vec{x}] \approx_\sigma N[\vec{P}/\vec{x}] \ .$$

**Remark 4.7.** Often in the literature, the observational equivalence is defined by considering multi-holed contexts. However, it is easy to see that the two notions of observational equivalences, obtained by considering just unary or all multi-holed contexts, coincide.

The problem of reducing the set of contexts in which we need to check the behavior of two terms has been widely studied in the literature. In particular, for both strategies in Definition 4.3 above, a *Context Lemma* holds, which allows us to restrict ourselves to *applicative contexts* of the shape $[\ ]\vec{P}$ ($[\ ]\vec{V}$), where $\vec{P}$ ($\vec{V}$) denotes a list of closed terms (values). Let us denote by $\approx_\sigma^{app}$ the observational equivalence which checks the behavior of terms only in applicative (by-value) contexts. This admits a coinductive characterization as follows:

**Definition 4.8** (Applicative $\sigma$-bisimilarity)**.**

(i) A relation $\mathcal{R} \subseteq \Lambda^0 \times \Lambda^0$ is

- an *applicative lazy bisimulation* if the following holds: $\langle M, N\rangle \in \mathcal{R} \implies (M \Downarrow_l \Leftrightarrow N \Downarrow_l) \wedge \forall P \in \Lambda^0.\ \langle MP, NP\rangle \in \mathcal{R}$.
- an *applicative cbv bisimulation* if the following holds: $\langle M, N\rangle \in \mathcal{R} \implies (M \Downarrow_v \Leftrightarrow N \Downarrow_v) \wedge \forall V$ closed value. $\langle MV, NV\rangle \in \mathcal{R}$.

(ii) The applicative equivalence $\approx_\sigma^{app}$ is the largest applicative bisimulation.

The following is a well-known result [AO93, EHR92]:

**Lemma 4.9** (Context Lemma)**.** $\approx_\sigma = \approx_\sigma^{app}$.

By the Context Lemma, the class of contexts in which we have to check the behavior of terms is smaller, however it is still infinite, thus the applicative bisimilarity is infinitely branching. In the following, we will study alternative coinductive characterizations of the observational equivalences, arising from the application of Leifer-Milner technique.

4.2. **Lambda Calculus as a Reactive System.** Both lazy and cbv $\lambda$-calculus can be endowed with a structure of reactive system in the sense of Definition 2.1, by considering corresponding context categories.

**Definition 4.10** (Lazy, cbv $\lambda$-reactive Systems)**.** $\mathbf{C}_\sigma^\lambda$, for $\sigma \in \{l, v\}$, consists of

- the category whose objects are $0, 1$, where the morphisms from 0 to 1 are the closed terms (up-to $\alpha$-equivalence), the morphisms from 1 to 1 are the unary closed contexts (up-to $\alpha$-equivalence), and composition is context insertion;
- the subcategory of reactive contexts is determined by the reactive contexts for the lazy and cbv strategy, respectively, presented in Remark 4.5;
- the (infinitely many) reaction rules are $(\lambda x.M)N \to_{\beta_\sigma} M[N/x]$, for all $M, N$.

The above definition is well-posed, in particular the subcategory of reactive contexts is composition-reflecting.

One can easily check that the reactive system $\mathbf{C}_\sigma^\lambda$ has redex RPOs; this fact can be proved by rephrasing the corresponding proof for the category of term contexts of [Sew02]. Here it is essential the fact that we consider only closed terms and closed contexts.

**Lemma 4.11.** *The reactive system* $\mathbf{C}_\sigma^\lambda$*, for* $\sigma \in \{l, v\}$*, has redex RPOs.*

The IPO contexts of a closed term for the lazy and cbv reactive systems are summarized in the second columns of the tables in Fig. 3.

**Lazy IPO lts's**

| term | IPO contexts | reactive IPO contexts |
|---|---|---|
| $\lambda x.M$ | $[\,]P$, $(\lambda x.C[\,])P$, $PC[\,]$ | $[\,]$P |
| $(\lambda x.M)N\vec{P}$ | $[\,]$, $(\lambda x.C[\,])P$, $PC[\,]$ | $[\,]$ |

**Cbv IPO lts's**

| term $M$ | IPO contexts | reactive IPO contexts |
|---|---|---|
| $\lambda x.M_1$ | $[\,]P$, $(\lambda x.C[\,])P$, $RC[\,]$, $(\lambda x.Q)C_1[\,]$ | $[\,]P$, $(\lambda x.Q)[\,]$ |
| $(\lambda x.M_1)N\vec{P}$ | $[\,]$, $(\lambda x.C[\,])P$, $RC[\,]$, $(\lambda x.Q)C_1[\,]$ | $[\,]$ |

where $R$ is not a value and $C_1[M]$ is a value.

Figure 3: IPO contexts for the lazy/cbv lts's.

The strong versions of context and IPO bisimilarities are too fine, since they take into account reaction steps, and tell apart $\beta$-convertible terms. Trivially, $I$ and $II$, where $I = \lambda x.x$, are equivalent neither in the context bisimilarity nor in the IPO bisimilarity, since $I \stackrel{[\,]}{\not\to}$, while $II \stackrel{[\,]}{\to}$ (both in the lazy and cbv case). On the other hand, one can easily check that the weak context bisimilarity, where the identity context $[\,]$ is unobservable, equates all closed terms. The appropriate notion is that of weak IPO bisimilarity, which, as we will see, turns out to capture exactly the lazy and cbv equivalences.

It is interesting to observe that also the observational equivalence and the applicative bisimilarity can be characterized as weak bisimilarities on suitable context lts's. In fact it is easy to prove that the observational equivalence $\approx_\sigma$ coincides with the weak bisimilarity on a restriction of the context lts built on $\mathbf{C}^\lambda_\sigma$, defined by $M \stackrel{C[\,]}{\longrightarrow} N$ iff $M \stackrel{C[\,]}{\longrightarrow}_C N$ and $M \Downarrow_\sigma$. Similarly, the applicative equivalence can be characterized by considering only applicative contexts in the lts.

In the following we will show that all these lts's induce the same notion of equivalence. Moreover, using the results of Section 3, we will show that the set of IPO contexts in the weak IPO bisimilarity to be considered can be significantly simplified. Then, from the fact that the weak IPO bisimilarity is the smallest one, it follows that it induces the simplest proofs that two terms are bisimilar.

Now, let us denote by $\approx_{\sigma I}$, for $\sigma \in \{l, v\}$, the lazy/cbv *weak IPO bisimilarity*, where the identity context is unobservable. In order to prove that $\approx_{\sigma I}$ is a congruence w.r.t. all contexts, we need to consider the category $\mathcal{D}^\lambda_\sigma$, list extension of $\mathcal{C}^\lambda_\sigma$, where the objects are finite lists $\langle 1, \ldots, 1\rangle$, and an arrow from $\underbrace{\langle 1, \ldots, 1\rangle}_n \to \underbrace{\langle 1, \ldots, 1\rangle}_m$ is a m-tuple of possibly multi-holed contexts $\langle C_1, \ldots, C_m\rangle$ with n holes all together. Then, in the lazy case one can show that any multi-holed context either is IPO uniform or it is of the shape $[\,]C_1[\,] \ldots C_k[\,]$ with the first hole reactive. Similarly, for the cbv case, all the multi-holed contexts are IPO uniform, apart from the contexts ranging on the following grammar, which have a reactive hole:

$$D[\,] \;::=\; [\,] \mid D[\,]C[\,] \mid (\lambda x.C[\,])D[\,]\;,$$

where $C$ is a multi-holed context. Moreover, the reduction relation is obviously deterministic. Thus, by applying Proposition 3.5, we have:

**Corollary 4.12.**

(i) *For all* $M, N \in \Lambda^0$*, for any closed unary context* $C[\ ]$*,*

$$M \approx_{\sigma I} N \implies C[M] \approx_{\sigma I} C[N] \ .$$

(ii) *Moreover*

$$\approx_{\sigma I} = \approx_{\sigma R} \ ,$$

*where* $\approx_{\sigma R}$ *denotes the weak IPO bisimilarity where only reactive contexts are considered (see the third columns in the tables of Fig. 3).*

Now, we are left to prove that the IPO bisimilarity coincides with the original observational equivalence.

**Proposition 4.13.** $\approx_l = \approx_l^{app} = \approx_{lI}$ *and* $\approx_v = \approx_{vI}$*.*

*Proof.* For the lazy case, we proceed by proving the following chain of inclusions:

$$\approx_l \subseteq \approx_l^{app} \subseteq \approx_{lR} \subseteq \approx_{lI} \subseteq \approx_l \ . \tag{4.1}$$

The first inclusion in the above chain (4.1), $\approx_l \subseteq \approx_l^{app}$, holds by definition. The third inclusion, $\approx_{lR} \subseteq \approx_{lI}$, follows by Corollary 4.12(ii). The others are proved as follows:

- $\approx_l^{app} \subseteq \approx_{lR}$. One can easily check that $\approx_l^{app}$ is a "weak IPO reactive bisimulation", using the fact that $\approx_l^{app}$ is closed under $\beta$-reduction.
- $\approx_{lI} \subseteq \approx_l$. Let $M \approx_{lI} N$. We have to show that, for any unary closed context $C[\ ]$, $C[M] \Downarrow \Leftrightarrow C[N] \Downarrow$. From $M \approx_{lI} N$, by Corollary 4.12(i), we have $C[M] \approx_{lI} C[N]$. Thus, if $C[M] \Downarrow$, then also $C[N] \Downarrow$, otherwise they would have different IPO reductions.

Notice that we provide a new alternative proof of the Context Lemma.

For the cbv case, considering the applicative equivalence $\approx_v^{app}$ does not help, but one can prove directly:

$$\approx_v \subseteq \approx_{vR} \subseteq \approx_{vI} \subseteq \approx_v \tag{4.2}$$

- $\approx_v \subseteq \approx_{vR}$. One can easily check that $\approx_v$ is a "weak IPO reactive bisimulation", using the fact that $\approx_v$ is closed under $\beta$-reduction.
- $\approx_{vR} \subseteq \approx_{vI}$. Immediate by Corollary 4.12(ii).
- $\approx_{vI} \subseteq \approx_v$. Let $M \approx_{vI} N$. We have to show that, for any unary context $C[\ ]$, $C[M] \Downarrow \Longleftrightarrow C[N] \Downarrow$. From $M \approx_{vI} N$, by Corollary 4.12(i), we have $C[M] \approx_{vI} C[N]$. Thus, if $C[M] \Downarrow$, then also $C[N] \Downarrow$, otherwise they would have different IPO reductions.

□

**Remark 4.14.** Corollary 4.12(ii) allows us to reduce the set of IPO contexts to be considered in the IPO bisimilarities. For the lazy case, only applicative contexts can be considered (see the first table in Figure 3), while for the cbv case, the set of reactive IPO contexts is larger (see the second table in Figure 3). However, also for the cbv case, one can prove that applicative (by-value) IPO contexts are sufficient. We omit the details.

Proposition 4.13 above gives us interesting characterizations of lazy and cbv observational equivalences, in terms of lts's where the labels are significantly reduced. However, such lts's (and bisimilarities) are still infinitely branching, e.g. $\lambda x.M \xrightarrow{P}_I$, for all $P \in \Lambda^0$. This is due to the fact that the context categories underlying the reactive systems $\mathbf{C}_l^\lambda$ and $\mathbf{C}_v^\lambda$ allow only for a ground representation of the $\beta$-rule through infinitely many ground rules. In order to overcome this problem, one should look for alternative categories which allow for a parametric representation of the $\beta$-rule as $(\lambda x.X)Y \to X[Y/x]$, where $X, Y$ are parameters. To this aim, we introduce the category of *second-order term contexts* (see Section 6 below). However, as we will see, this approach works only if the reaction rules are "local", that is they do not act on the whole term, but only locally. In particular, the operation of substitution on the $\lambda$-calculus is not local and thus it is not describable by a finite set of reaction rules. To avoid this problem, in the following section we consider encodings of the $\lambda$-calculus into Combinatory Logic (CL) endowed with suitable strategies and equivalences, which turn out to correspond to lazy and cbv equivalences.

## 5. Combinatory Logic

In this section, we focus on *Combinatory Logic* [HS86] with Curry's combinators $\mathbf{K}, \mathbf{S}$, and we study its relationships with the $\lambda$-calculus endowed with lazy and cbv reduction strategies. An interesting result that we prove is that we can define suitable reduction strategies on CL-terms, inducing observational equivalences which correspond to lazy and cbv equivalences on $\lambda$-calculus. As a consequence, we can safely shift our attention from the reactive system of $\lambda$-calculus to the simpler reactive system of CL. In this section, we apply Leifer-Milner construction to CL viewed as a (standard) context category, and we study weak versions of context and IPO bisimilarities. Our main result is that we can recover lazy and cbv observational equivalences as weak IPO equivalences on CL$^*$, a variant of standard CL. Here the approach is first-order, thus the IPO equivalences are still infinitely branching. However, the results in this section are both interesting in themselves, and useful for our subsequent investigation of Section 6, where CL is viewed as a second-order rewriting system, and characterizations of the observational equivalences as finitely branching IPO bisimilarities are given.

**Definition 5.1** (Combinatory Terms)**.** The set of combinatory terms is defined by:

$$(CL \ni)\ M \ ::=\ x \mid \mathbf{K} \mid \mathbf{S} \mid MM\ ,$$

where $\mathbf{K}$, $\mathbf{S}$ are combinators.
Let $CL^0$ denote the set of *closed* CL-terms.

### 5.1. Correspondence with the $\lambda$-calculus.

Let $\Lambda(\mathbf{K}, \mathbf{S})$ denote the set of $\lambda$-terms built over constants $\mathbf{K}, \mathbf{S}$. The following is a well-known encoding:

**Definition 5.2** ($\lambda$-encoding)**.** Let $\mathcal{T} : \Lambda(\mathbf{K}, \mathbf{S}) \to CL$ be the transformation defined as follows:

$$\begin{array}{ll}
\mathcal{T}(x) = x & \mathcal{T}(C) = C \ \text{ if } C \in \{\mathbf{K}, \mathbf{S}\} \\
\mathcal{T}(MN) = \mathcal{T}(M)\mathcal{T}(N) & \mathcal{T}(\lambda x.MN) = \mathbf{S}\mathcal{T}(\lambda x.M)\mathcal{T}(\lambda x.N) \\
\mathcal{T}(\lambda x.x) = \mathbf{SKK} & \mathcal{T}(\lambda x.\lambda y.M) = \mathcal{T}(\lambda x.\mathcal{T}(\lambda y.M)) \\
\mathcal{T}(\lambda x.y) = \mathbf{K}y & \mathcal{T}(\lambda x.C) = \mathbf{K}\mathcal{T}(C) \ \text{ if } C \in \{\mathbf{K}, \mathbf{S}\}
\end{array}$$

In particular, if we restrict the domain of $\mathcal{T}$ to $\Lambda$, we get an encoding of the $\lambda$-calculus into CL.
Vice versa, there is a natural embedding of CL into the $\lambda$-calculus $\mathcal{E} : CL \to \Lambda$:

$$\mathcal{E}(\mathbf{K}) = \lambda xy.x \qquad \mathcal{E}(\mathbf{S}) = \lambda xyz.(xz)(yz) \qquad \mathcal{E}(x) = x \qquad \mathcal{E}(MN) = \mathcal{E}(M)\mathcal{E}(N)$$

The following lemma holds:

**Lemma 5.3.** *For all $M \in \Lambda$, $\mathcal{E}(\mathcal{T}(M)) =_\sigma M$, for $\sigma \in \{\beta, \beta_V\}$.*

*Proof.* First, one can easily prove that, if $M$ is $\lambda$-free, then $\mathcal{E}\mathcal{T}(\lambda x.M) =_\sigma \lambda x.M$ (by induction on $M$). Then, using the fact that $\mathcal{T}(M)$ is $\lambda$-free for all $M$, by definition of $\mathcal{T}$, one gets that $\mathcal{T}^2(M) = \mathcal{T}(M)$ for all $M$. Finally, we are ready to prove the thesis in its full generality by induction on $M$. The only non-trivial case is when $M = \lambda x.\lambda y.N$. Then we have $\mathcal{E}\mathcal{T}(\lambda x.\lambda y.N) = \mathcal{E}\mathcal{T}(\lambda x.\mathcal{T}(\lambda y.N))$, where $\mathcal{T}(\lambda y.N) = PQ$ is $\lambda$-free. Then
$\mathcal{E}\mathcal{T}(\lambda x.\lambda y.N) = \mathcal{E}\mathcal{T}(\lambda x.PQ) = \mathcal{E}(S)\mathcal{E}\mathcal{T}(\lambda x.P)\mathcal{E}\mathcal{T}(\lambda x.Q) = S\mathcal{E}\mathcal{T}(\lambda x.P)\mathcal{E}\mathcal{T}(\lambda x.Q)$
$=_\sigma S\lambda x.P\lambda x.Q$, since $PQ$ is $\lambda$-free,
$=_\sigma \lambda x.PQ =_\sigma \lambda x.\mathcal{E}\mathcal{T}(PQ)$, since $PQ$ is $\lambda$-free,
$= \lambda x.\mathcal{E}\mathcal{T}\mathcal{T}(\lambda y.N) = \lambda x.\mathcal{E}\mathcal{T}(\lambda y.N)$, since $\mathcal{T}^2 = \mathcal{T}$,
$=_\sigma \lambda x.\lambda y.N$, by induction hypothesis. □

5.1.1. *Lazy/cbv observational equivalence on CL.* Usually, the set of combinatory terms are endowed with the following reaction rules:

$$\mathbf{K}MN \to M \qquad\qquad \mathbf{S}MNP \to (MP)(NP)$$

We will also consider a cbv version of the above rules, reducing CL redexes only when the arguments are values, i.e. terms on the following grammar:

$$V \ ::= \ \mathbf{K} \mid \mathbf{S} \mid \mathbf{K}V \mid \mathbf{S}V \mid \mathbf{S}VV \ .$$

The cbv rules are the following:

$$\mathbf{K}V_1V_2 \to V_1 \qquad\qquad \mathbf{S}V_1V_2V_3 \to (V_1V_3)(V_2V_3)$$

**Definition 5.4** (Lazy/cbv Reduction Strategy on CL)**.**

(i) The *lazy reduction strategy* $\to_l \subseteq CL^0 \times CL^0$ reduces the leftmost outermost CL-redex. Formally:

$$\frac{}{\mathbf{S}M_1M_2M_3 \to_l (M_1M_3)(M_2M_3)} \qquad \frac{}{\mathbf{K}M_1M_2 \to_l M_1} \qquad \frac{M \to_l M'}{MP \to_l M'P}$$

(ii) The *cbv strategy* $\to_v \subseteq CL^0 \times CL^0$ is defined by

$$\frac{}{\mathbf{S}V_1V_2V_3 \to_v (V_1V_3)(V_2V_3)} \qquad \frac{}{\mathbf{K}V_1V_2 \to_v V_1} \qquad \frac{M_1 \to_v M_1'}{\mathbf{K}M_1 \to_v \mathbf{K}M_1'}$$

$$\frac{M_2 \to_v M_2'}{\mathbf{K}V_1M_2 \to_v \mathbf{K}V_1M_2'} \qquad \frac{M_1 \to_v M_1'}{\mathbf{S}M_1 \to_v \mathbf{S}M_1'} \qquad \frac{M_2 \to_v M_2'}{\mathbf{S}V_1M_2 \to_v \mathbf{S}V_1M_2'}$$

$$\frac{M_3 \to_v M_3'}{\mathbf{S}V_1V_2M_3 \to_v \mathbf{S}V_1V_2M_3'} \qquad \frac{M \to_v M'}{MP \to_v M'P}$$

where $V_1, V_2, V_3$ are values.

**Definition 5.5** (Unary Contexts on CL)**.** The set of *unary contexts* on CL is defined by

$$C[\,] \ ::= \ [\,] \mid C[\,]P \mid PC[\,] \ .$$

Alternatively we could define the lazy strategy $\to_l$ as the closure of the standard CL-reaction rules under the following reactive contexts (which coincide with the applicative ones):

$$D[\,] \ ::= \ [\,] \mid D[\,]P \ .$$

Similarly, by considering the restriction to values of the reaction rules of Definition 5.1, we could define the cbv strategy $\to_v$ as the closure of the cbv reaction rules under the following reactive contexts:

$$D[\,] \ ::= \ [\,] \mid D[\,]P \mid \mathbf{K}D[\,] \mid \mathbf{K}VD[\,] \mid \mathbf{S}D[\,] \mid \mathbf{S}VD[\,] \mid \mathbf{S}V_1V_2D[\,].$$

Let $\downarrow_\sigma$ denote the convergence relation on CL, for $\sigma \in \{l, v\}$.

**Definition 5.6** (Lazy/cbv Equivalence on CL)**.**

(i) A relation $\mathcal{R} \subseteq CL^0 \times CL^0$ is a
- *CL lazy bisimulation* if:
  $\langle M, N\rangle \in \mathcal{R} \implies (M \downarrow_l \Leftrightarrow N \downarrow_l) \wedge \forall P \in \Lambda^0.\ \langle MP, NP\rangle \in \mathcal{R}$.
- *CL cbv bisimulation* if:
  $\langle M, N\rangle \in \mathcal{R} \implies (M \downarrow_v \Leftrightarrow N \downarrow_v) \wedge \forall$ closed value $V \in \Lambda^0.\ \langle MV, NV\rangle \in \mathcal{R}$.

(ii) Let $\simeq_\sigma \subseteq CL^0 \times CL^0$ be the largest CL lazy/cbv bisimulation.

(iii) Let $\widehat{\simeq}_\sigma \subseteq CL \times CL$ denote the extension of $\simeq_\sigma$ to open terms defined by: for $M, N \in CL$ s.t. $FV(M, N) \subseteq \{x_1, \dots, x_n\}$, $M \widehat{\simeq}_\sigma N$ iff for all closing (by-value) substitutions $\vec{P}$, $M[\vec{P}/\vec{x}] \simeq_\sigma N[\vec{P}/\vec{x}]$.

Notice that we use two different symbols for equivalences ($\approx$ and $\simeq$), in this way we distinguish the equivalence relation on $\lambda$-terms from the corresponding relation on CL.

The following theorem is interesting *per se*:

**Theorem 5.7.** *For all $M, N \in \Lambda$, $M \widehat{\approx}_\sigma N \iff \mathcal{T}(M) \widehat{\simeq}_\sigma \mathcal{T}(N)$ .*

**Proof of Theorem 5.7**. We carry out the proof of the above theorem for the lazy case, the proof for the cbv case being similar.

**Lemma 5.8.**

(i) *For all $M \in CL^0$, $M \downarrow_l \iff \mathcal{E}(M) \Downarrow_l$.*

(ii) *For all $M \in \Lambda^0$, $M \Downarrow_l \iff \mathcal{T}(M) \downarrow_l$.*

*Proof.* (i) By definition of the lazy strategies on $\lambda$-terms and on CL-terms.
(ii) ($\Rightarrow$) Let $M \Downarrow_l$. Then, since by Lemma 5.3 $\mathcal{E}(\mathcal{T}(M)) =_\beta M$, and $\approx_l$ is closed under $\beta$-conversion, we have also $\mathcal{E}(\mathcal{T}(M)) \Downarrow_l$. Thus, by (i), $\mathcal{T}(M) \downarrow_l$.
(ii) ($\Leftarrow$) Let $\mathcal{T}(M) \downarrow_l$. By (i), $\mathcal{E}(\mathcal{T}(M)) \Downarrow_l$, by Lemma 5.3, $M =_\beta \mathcal{E}(\mathcal{T}(M))$, thus $M \Downarrow_l$. □

**Lemma 5.9.** *For all $M, N \in CL^0$, if $\mathcal{E}(M) =_\beta \mathcal{E}(N)$, then $M \simeq_l N$.*

*Proof.* The proof follows from the fact that $\mathcal{R} = \{\langle M, N\rangle \in CL^0 \mid \mathcal{E}(M) =_\beta \mathcal{E}(N)\}$ is a CL lazy bisimulation. Namely $M \downarrow_l$ iff $N \downarrow_l$, because, by Lemma 5.8(i), $M \downarrow_l$ iff $\mathcal{E}(M) \downarrow_l$ and $N \downarrow_l$ iff $\mathcal{E}(N) \downarrow_l$, and $\approx_l^{app}$ is closed under $\beta$-conversion. Moreover, for any $P \in CL^0$, $\langle MP, NP\rangle \in \mathcal{R}$, since $\mathcal{E}(MP) = \mathcal{E}(M)\mathcal{E}(P) =_\beta \mathcal{E}(N)\mathcal{E}(P) = \mathcal{E}(NP)$. □

**Lemma 5.10.** $\forall P \in CL^0$*, $P \simeq_l \mathcal{T}(\mathcal{E}(P))$.*

*Proof.* We prove that $\mathcal{R} = \{(P\vec{R}, \mathcal{T}(\mathcal{E}(P))\vec{R}) \mid P, \vec{R} \in CL^0\}$ is a bisimulation. To this aim, it is sufficient to prove that, for all $P, \vec{R}$, $P\vec{R} \downarrow_l \Leftrightarrow \mathcal{T}(\mathcal{E}(P))\vec{R} \Downarrow_l$. By Lemma 5.8, $P\vec{R} \downarrow_l \Leftrightarrow \mathcal{E}(P\vec{R}) \Downarrow_l$. Now $\mathcal{E}(P\vec{R}) = \mathcal{E}(P)\mathcal{E}(\vec{R}) =_\beta (\mathcal{E} \circ \mathcal{T} \circ \mathcal{E}(P))\mathcal{E}(\vec{R}) = \mathcal{E}((\mathcal{T} \circ \mathcal{E}(P))\vec{R})$. Finally, by Lemma 5.8, $\mathcal{E}((\mathcal{T} \circ \mathcal{E}(P))\vec{R} \Downarrow_l \iff \mathcal{T}(\mathcal{E}(P))\vec{R} \downarrow_l$. □

5.2. **The First-order Approach: CL as a Context Category.** We endow CL with a structure of reactive system in the sense of [LM00], by considering the context category of closed unary contexts:

**Definition 5.11** (Lazy, cbv CL Reactive Systems)**.** $\mathbf{C}^1_\sigma$, for $\sigma \in \{l, v\}$, consists of:

- the context category whose objects are $0, 1$, where the morphisms from 0 to 1 are the closed terms, the morphisms from 1 to 1 are the closed unary contexts, and composition is context substitution;
- the subcategory of reactive contexts is determined by the reactive contexts for the lazy and cbv strategy, respectively, presented in Definition 5.4;
- the reaction rules are the standard CL reduction rules for the lazy case, and the cbv reduction rules for the cbv case.

**Lemma 5.12.** *The reactive systems $\mathbf{C}^1_\sigma$ have redex RPOs.*

One can easily check that the IPO contexts are the following.

- **Lazy**. The IPO contexts for a given term $M$ are:
  - $[\,]\vec{P}$, where $\vec{P}$ has the minimal length for the top-level reaction of $M$ to fire,
  - $\mathbf{K}C[\,]P_1,\ \mathbf{K}P_1C[\,],\ \mathbf{K}P_1\vec{Q}C[\,]$, for any $C[\,], \vec{Q}, P_1$,
  - $\mathbf{S}C[\,]P_1P_2,\ \mathbf{S}P_1C[\,]P_2,\ \mathbf{S}P_1P_2C[\,],\ \mathbf{S}P_1P_2\vec{Q}C[\,]$, for any $P_1, P_2, C[\,], \vec{Q}$.
- **Cbv**.
  For $M$ not a value, the following contexts are IPOs:
  - $[\,]$,

  For $M$ value, the following contexts are IPOs:
  - $[\,]V_1 \dots V_i$, where $i$ is the minimum number of arguments necessary for the top-level reaction of $M$ to fire,
  - $[\,]V_1 \dots V_iP$, where $P$ is not a value, and $i$, possibly 0, is less than the minimum number of arguments necessary for the top-level reaction of $M$ to fire,
  - $VC[\,]V_1 \dots V_i$ where $V$ and $C[M]$ are values and $i+1$ is the minimum number of arguments necessary for the top-level reaction of $V$ to fire, in more detail: $\mathbf{K}C[\,]V$, $\mathbf{K}VC[\,]$, $\mathbf{S}C[\,]V_1V_2$, $\mathbf{S}V_1C[\,]V_2$, $\mathbf{S}V_1V_2C[\,]$,
  - $VC[\,]V_1 \dots V_iP$, where $V$ and $C[M]$ are values, $P$ is not a value, and $i+1$ is less than the minimum number of arguments necessary for the top-level reaction of $V$ to fire, in more detail: $\mathbf{K}C[\,]P$, $\mathbf{S}C[\,]P$, $\mathbf{S}C[\,]V_1P$, $\mathbf{S}V_1C[\,]P$.

  For any term $M$, the following contexts are IPOs:
  - $PC[\,]$, where $P$ is not a value and $C[\,]$ is any context.

The strong versions of context and IPO bisimilarities are too fine, since, as in the $\lambda$-calculus case, they take into account reduction steps, and tell apart $\beta$-convertible terms. Thus we consider weak variants of such equivalences, where the identity context $[\,]$ is unobservable. Weak context bisimilarity is too coarse, since it equates all terms. However, we will prove that the weak IPO bisimilarity "almost" coincides with the lazy/cbv equivalence. Moreover, we will show how to recover the exact correspondence by considering a suitable variant of CL.

First of all, let $\simeq_{\sigma I}$, for $\sigma \in \{l, v\}$, denote the *lazy/cbv weak IPO bisimilarity* obtained by considering the identity context as unobservable. Similarly to the case of the $\lambda$-calculus, we can define a list extension categories by taking the category of multi-holed contexts. In this category all contexts with no reactive indexes are IPO uniform. In fact, given a multi-holed context $C[\ ,\dots,\ ]$, consider the outermost, leftmost hole in $C[\ ,\dots,\ ]$, if this hole lies in a reactive position, that is if the context obtained by replacing the other holes with a given term, $\mathbf{K}$ for example, is reactive, then $C[\ ,\dots,\ ]$ has a reactive index, otherwise is uniform. For the lazy case the multi-holed contexts with a reactive index are given by the grammar:

$$D[\,] ::= [\,] \mid D[\,]C[\,] \,.$$

For the cbv case:

$$D[\,]\ ::=\ [\,] \mid D[\,]C[\,] \mid \mathbf{K}D[\,] \mid \mathbf{K}VD[\,] \mid \mathbf{S}D[\,] \mid \mathbf{S}VD[\,] \mid \mathbf{S}V_1V_2D[\,].$$

Thus, by Proposition 3.5(i), we have:

**Proposition 5.13.** *For all $M, N \in CL^0$, for any closed unary context $C[\,]$,*

$$M \simeq_{\sigma I} N \implies C[M] \simeq_{\sigma I} C[N]\ .$$

The rest of this section is devoted to compare the lazy/cbv weak IPO bisimilarity $\simeq_{\sigma I}$ with the lazy/cbv equivalence on CL $\simeq_\sigma$ defined in Definition 5.6. The following lemma can be easily proved by coinduction, using Proposition 5.13.

**Lemma 5.14.** $\simeq_{\sigma I} \subseteq \simeq_\sigma$.

*Proof.* We prove that $\simeq_{\sigma I}$ is a lazy/cbv bisimulation on CL. Let $M \simeq_{\sigma I} N$. If $M \downarrow_\sigma$, then also $N \downarrow_\sigma$, since a convergent term has different IPO-transitions from a divergent term. We are left to prove that for all $P$, $MP \simeq_{\sigma I} NP$. But this follows from Proposition 5.13. □

However, the converse inclusion $\simeq_\sigma \subseteq \simeq_{\sigma I}$ does not hold, since for instance $\mathbf{K} \simeq_\sigma \mathbf{S}(\mathbf{KK})(\mathbf{SKK})$, because, e.g. for the lazy case, for all $P$, $\mathbf{S}(\mathbf{KK})(\mathbf{SKK})P \to^* \mathbf{K}P$. But $\mathbf{K} \not\simeq_{\sigma I} \mathbf{S}(\mathbf{KK})(\mathbf{SKK})$. Namely $\mathbf{S}(\mathbf{KK})(\mathbf{SKK}) \xrightarrow{[\,]V}_I$, while $\mathbf{K} \not\xrightarrow{[\,]V}_I$. The problem arises since the equivalence $\simeq_{\sigma I}$ tells apart terms whose top-level combinators expect a different number of arguments to reduce. In order to overcome this problem, we consider an extended calculus, CL$^*$, where the combinators $\mathbf{K}$ and $\mathbf{S}$ become unary, at the price of adding new intermediate combinators and intermediate reductions (the reactive contexts are the ones in Definition 5.11).

**Definition 5.15.** The CL$^*$ lazy combinatory calculus is defined by

- Terms: $M\ ::=\ x \mid \mathbf{K} \mid \mathbf{S} \mid \mathbf{K}'M \mid \mathbf{S}'M \mid \mathbf{S}''MN \mid MN$
  where $\mathbf{K}$, $\mathbf{K}'$, $\mathbf{S}$, $\mathbf{S}'$, $\mathbf{S}''$ are combinators.
- Rules: $\mathbf{K}M \to \mathbf{K}'M$ $\quad$ $\mathbf{K}'MN \to M$
  $\mathbf{S}M \to \mathbf{S}'M$ $\quad$ $\mathbf{S}'MN \to \mathbf{S}''MN$ $\quad$ $\mathbf{S}''MNP \to (MP)(NP)$

The CL$^*$ cbv combinatory calculus is defined by

- Terms: $M\ ::=\ x \mid \mathbf{K} \mid \mathbf{S} \mid MN \mid \mathbf{K}'V \mid \mathbf{S}'V \mid \mathbf{S}''VV$
  Values: $V\ ::= \mathbf{K} \mid \mathbf{K}'V \mid \mathbf{S} \mid \mathbf{S}'V \mid \mathbf{S}''VV$
  where $\mathbf{K}$, $\mathbf{K}'$, $\mathbf{S}$, $\mathbf{S}'$, $\mathbf{S}''$ are combinators.
- Rules: $\mathbf{K}V_1 \to \mathbf{K}'V_1$ $\quad$ $\mathbf{K}'V_1V_2 \to V_1$
  $\mathbf{S}V_1 \to \mathbf{S}'V_1$ $\quad$ $\mathbf{S}'V_1V_2 \to \mathbf{S}''V_1V_2$ $\quad$ $\mathbf{S}''V_1V_2V_3 \to (V_1V_3)(V_2V_3)$

Notice that the calculus in the above definition is well-defined, since the set of terms is closed under the reaction rules. One can define lazy/cbv reduction strategies on CL$^*$ as in Definition 5.4, or as the closures of the reaction rules under the following reactive contexts:

**Definition 5.16** (CL$^*$ Reactive Contexts)**.**

- **Lazy**. $D[\,]\ ::=\ [\,] \mid D[\,]P$ .
- **Cbv**. $D[\,]\ ::=\ [\,] \mid D[\,]P \mid VD[\,]$.

Let $\simeq^*_\sigma$ be the lazy/cbv equivalence defined on CL$^*$, similarly as in Definition 5.6 for CL. There is a trivial embedding of CL-terms into CL$^*$. Moreover, one can easily check that, when restricted to terms of CL, $\simeq^*_\sigma$ coincides with $\simeq_\sigma$.

**Lazy IPO lts's on CL$^*$**

| term $M$ | IPO contexts | reactive IPO contexts |
|---|---|---|
| $M$ value | $[\,]P$, $PC[\,]$ | $[\,]P$ |
| $M$ not a value | $[\,]$, $PC[\,]$ | $[\,]$ |

**Cbv IPO lts's on CL$^*$**

| term $M$ | IPO contexts | reactive IPO contexts |
|---|---|---|
| $M$ value | $[\,]P$, $RC[\,]$, $V[\,]$ | $[\,]P$, $V[\,]$ |
| $M$ not a value | $[\,]$, $RC[\,]$ | $[\,]$ |

where $R$ is not a value, $V$ is a value, $C[\,]$ is a generic unary context.

Figure 4: IPO contexts for the lazy/cbv lts's on CL$^*$.

Analogously to the CL case, we define the reactive system over CL$^*$. In the context category, the unary closed contexts are defined by the grammar

$$C[\,] \ ::= \ [\,] \mid C[\,]M \mid MC[\,]$$

where $M$ is a closed term. Notice that, under the above definition, expressions like $\mathbf{K}'[\,]$ do not represent unary closed context. In defining the IPO transitions, it is important to observe that $C[M]$ is a value iff $M$ is a value and $C[\,]$ is the identity context $[\,]$. Let us denote by $\simeq^*_{\sigma I}$ the weak IPO bisimilarity obtained by considering the lazy/cbv reactive system over CL$^*$. Since CL$^*$-terms expect at most one argument, the IPO contexts for CL$^*$ are simpler than the ones for CL, and they are summarized in Figure 4.

Similarly to the previous case, one can consider the multi-holed contexts category as a list extension category. In this category all contexts are either IPO uniform or have a reactive index. Moreover the reduction relation is deterministic. Thus Proposition 3.5 applies and we have:

**Proposition 5.17.**

(i) *The equivalence $\simeq^*_{\sigma I}$ is a congruence w.r.t. unary contexts.*

(ii) $\simeq^*_{\sigma I} = \simeq^*_{\sigma R}$, *where $\simeq^*_{\sigma R}$ denotes the IPO bisimilarity where only reactive IPO contexts are considered.*

By Proposition 5.17(ii) above, the weak IPO equivalence can be significantly simplified. Namely, in the lazy case, we obtain the weak IPO bisimilarity $\simeq_{lR}$, where only applicative IPO contexts are considered (see Figure 4). In the cbv case, Proposition 5.17 allows us to reduce ourselves to contexts of the shape $[\,], [\,]P, V[\,]$ (see Figure 4). However, one can prove that also in this case we can consider only applicative by-value contexts. We skip the details of such proof.

Moreover, we have $\mathbf{K} \simeq^*_{\sigma I} \mathbf{S}(\mathbf{KK})(\mathbf{SKK})$. More in general, the weak IPO bisimilarity $\simeq^*_{\sigma I}$ coincides with the lazy/cbv equivalence on CL:

**Theorem 5.18.** *For all $M, N \in CL^0$, $\quad M \simeq^*_{\sigma I} N \iff M \simeq_\sigma N$.*

*Proof.* ($\subseteq$) One can show that $\simeq^*_{\sigma I} \subseteq \simeq^*_\sigma$ by coinduction, as in the proof of Lemma 5.14, by showing that $\simeq^*_{\sigma I}$ is a bisimulation on CL$^*$, also using Proposition 5.17. Then, since $\simeq^*_\sigma$ coincides with $\simeq_\sigma$ on CL-terms, we have the thesis.
($\supseteq$) By coinduction, showing that $\simeq_\sigma$ is a weak IPO bisimulation on CL$^*$. $\square$

As a consequence of Theorem 5.7 and Theorem 5.18 above, we can recover the lazy/cbv observational equivalence on $\lambda$-terms as weak IPO bisimilarity on CL$^*$.

**Proposition 5.19.** *For all $M, N \in \Lambda^0$,* $\quad M \approx_\sigma N \iff \mathcal{T}(M) \simeq^*_{\sigma I} \mathcal{T}(N)$.

However, such notions of weak IPO bisimilarities still suffer of the problem of being infinitely branching, since the IPO contexts are [ ], [ ]$P$ for the lazy case, and [ ], [ ]$V$ for the cbv case, for all $P, V \in (CL^*)^0$.

This problem will be solved in the next section, where we introduce the notion of second-order context category, and we endow CL$^*$ with such a structure.

## 6. Second-order Term Contexts

The definition of term context category [LM00] can be generalized to a definition of second-order term context category. The generalization is obtained by extending the term syntax with function (second-order) variables, that is variables not standing for terms but instead for functions on terms. The formal definition is the following

**Definition 6.1** (Category of Second-order Term Contexts)**.** Let $\Sigma$ be a signature for a term language. The category of *second-order term contexts* over $\Sigma$ is defined by: objects are finite lists of naturals $\langle n_1, \ldots, n_k\rangle$, an arrow $\langle m_1, \ldots, m_h\rangle \to \langle n_1, \ldots, n_k\rangle$ is a k-tuple $\langle t_1, \ldots, t_k\rangle$, where the term $t_i$ is defined over the signature $\Sigma \cup \{F_1^{m_1}, \ldots, F_h^{m_h}\} \cup \{X_{i,1} \ldots, X_{i,n_i}\}$, where $F_i^{m_i}$ is a function variable of arity $m_i$, $X_{i,j}$ is a ground variable. The category of *second-order linear term context*, $T_2^*(\Sigma)$, is the subcategory whose arrows are n-tuples of terms, satisfying the condition that the n-tuples have to contain exactly one use of each function variable $F_i^{m_i}$.

One can check that the standard category of term contexts over $\Sigma$ coincides with the subcategory whose objects are the lists containing only copies of the natural number 0; in fact this subcategory uses function variables with no arguments and the ground variables do not appear.

In order to define composition in category $T_2^*(\Sigma)$, and more generally in the category of second-order term context, it is useful to represent morphisms, i.e. terms on the signature $\Sigma \cup \{F_1^{m_1}, \ldots, F_h^{m_h}\} \cup \{X_{i,1}, \ldots, X_{i,n_i}\}$, using a $\lambda$-notation for binding variables, that is, instead of writing a term with free variables, we write its lambda closure. In this way it is possible to define composition of arrows in terms $\beta$ reduction. To avoid ambiguities we use a different symbol $\boldsymbol{\lambda}$ for this later form of lambda abstraction. With this notation, a term $t$ on the signature $\Sigma \cup \{F_1^{m_1}, \ldots, F_h^{m_h}\} \cup \{X_1, \ldots, X_n\}$ is written as: $\boldsymbol{\lambda} F_1^{m_1} \ldots F_h^{m_h}.\boldsymbol{\lambda} X_1 \ldots X_n.t$, or as $\boldsymbol{\lambda}\vec{F}.\boldsymbol{\lambda}\vec{X}.t$ for brevity. In general a second-order context $\langle t_1, \ldots, t_k\rangle : \langle m_1, \ldots, m_h\rangle \to \langle n_1, \ldots, n_k\rangle$ is written as $\boldsymbol{\lambda}\vec{F}.\langle\boldsymbol{\lambda}\vec{X}_1.t_1, \ldots, \boldsymbol{\lambda}\vec{X}_k.t_k\rangle$.

- The identity on $\langle n_1, \ldots, n_k\rangle$ is: $\boldsymbol{\lambda}\vec{F}.\langle\boldsymbol{\lambda}\vec{X}_1.F_1^{n_1}(\vec{X}_1), \ldots, \boldsymbol{\lambda}\vec{X}_k.F_k^{n_k}(\vec{X}_k)\rangle$.
- The composition between the morphisms
  $\boldsymbol{\lambda}\vec{F}.\langle\boldsymbol{\lambda}\vec{X}_1.s_1, \ldots, \boldsymbol{\lambda}\vec{X}_k.s_k\rangle : \langle l_1, \ldots, l_h\rangle \to \langle m_1, \ldots, m_k\rangle$ and
  $\boldsymbol{\lambda}\vec{G}.\langle\boldsymbol{\lambda}\vec{Y}_1.t_1, \ldots, \boldsymbol{\lambda}\vec{Y}_j.t_j\rangle : \langle m_1, \ldots, m_k\rangle \to \langle n_1, \ldots, n_j\rangle$
  is the $\beta$-normal form of the expression
  $\boldsymbol{\lambda}\vec{F}.(\boldsymbol{\lambda}\vec{G}.\langle\boldsymbol{\lambda}\vec{Y}_1.t_1, \ldots, \boldsymbol{\lambda}\vec{Y}_j.t_j\rangle)\ (\boldsymbol{\lambda}\vec{X}_1.s_1, \ldots, \boldsymbol{\lambda}\vec{X}_k.s_k) : \langle l_1, \ldots, l_h\rangle \to \langle n_1, \ldots, n_j\rangle$.

Informally, the composition is given by a $j$-tuple of expressions $t_i$ in which every function variable $G_l$ is substituted by the corresponding expression $s_l$, with the ground variables of

$s_l$ substituted by the corresponding parameters of $G_l$ in $t_i$. For example, considering the signature for natural numbers $\{0, S, +\}$, the composition between $\boldsymbol{\lambda} F.\boldsymbol{\lambda} X_1.F(X_1, 0) : \langle 2\rangle \to \langle 1\rangle$ and $\boldsymbol{\lambda} G.\boldsymbol{\lambda} Y_1 Y_2.(G(S(Y_1)) + Y_2) : \langle 1\rangle \to \langle 2\rangle$ is the second order context $\boldsymbol{\lambda} F.\boldsymbol{\lambda} Y_1 Y_2.F(S(Y_1), 0) + Y_2 : \langle 2\rangle \to \langle 2\rangle$.

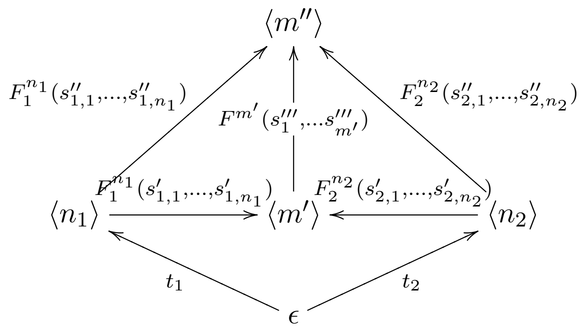
Note that the identity morphism is defined as a $\lambda$-term implementing the identity function, while composition on morphisms is defined by the function composition in the $\lambda$-setting. Given this correspondence, it is easy to prove that the categorical properties for the identity hold, while the associativity of composition essentially follows from the unicity of the normal form.

The main general result on second-order term contexts is the following:

**Proposition 6.2.** *For any signature $\Sigma$, in the category of second-order linear term contexts over $\Sigma$, any commuting square, having as initial vertex the empty list $\epsilon$, has an RPO.*

*Proof.* Here we present the proof just for the special case that it is useful in this paper, namely we consider the restricted category containing as objects the lists with at most one element. The proof for the complete category follows a similar pattern. Given two arrows with domain the empty list: $t_1 : \epsilon \to \langle n_1\rangle$ and $t_2 : \epsilon \to \langle n_2\rangle$, and two arrows $s_1 : \langle n_1\rangle \to \langle m\rangle$, $s_2 : \langle n_2\rangle \to \langle m\rangle$ forming a commuting square ($s_1 \circ t_1 = s_1 \circ t_1 : \epsilon \to \langle m\rangle$), the corresponding RPO for this commuting square is inductively defined on the structures of $s_1$, $s_2$. There are several cases to consider:
(i) $s_1 = c_1(s_{1,1}, \ldots, s_{1,k_1})$ and $s_2 = c_2(s_{2,1}, \ldots, s_{2,k_2})$, with $c_1, c_2$ function symbols in the signature $\Sigma$. Necessarily $c_1 = c_2$ (and $k_1 = k_2$). We have to consider in which subterms of $s_1$ and $s_2$ the function variables, $F_1^{n_1}$ and $F_2^{n_2}$, appear. If $F_1^{n_1}$ and $F_2^{n_2}$ appear in corresponding subterms, that is, there is an $i$ such that $F_1^{n_1}$ appears in $s_{1,i}$ and $F_2^{n_2}$ in $s_{2,i}$, then we have that $s_{1,i}$ and $s_{2,i}$, together with $t_1, t_2$, form a commuting square, and the RPO, inductively defined, for this second commuting square, immediately induces the RPO for $s_1$ and $s_2$. The subcase where $F_1^{n_1}$ and $F_2^{n_2}$ do not appear in corresponding subterms is treated at point (iii).
(ii) $s_1 = F_1^{n_1}(s_{1,1}, \ldots, s_{1,n_1})$ and $s_2 = F_2^{n_2}(s_{2,1}, \ldots, s_{2,n_2})$.
In this case, we have that $t_1[s_{1_1}/X_{1,1}, \ldots, s_{1,n_1}/X_{1,n_1}] = t_2[s_{2_1}/X_{2,1}, \ldots, s_{2,n_2}/X_{2,n_2}]$, that is there is a unifier i.e. a substitution making $t_1$ and $t_2$ equal. Consider the most general unifier (mgu) for $t_1$ and $t_2$, this is given by two tuples of terms, $s'_{1,1}, \ldots, s'_{1,n_1}$ and $s'_{2,1}, \ldots, s'_{2,n_2}$, such that $t_1[s'_{1_1}/X_{1,1}, \ldots, s'_{1,n_1}/X_{1,n_1}] = t_2[s'_{2_1}/X_{2,1}, \ldots, s'_{2,n_2}/X_{2,n_2}]$.
We have that: $F_1^{n_1}(s'_{1,1}, \ldots, s'_{1,n_1}) : \langle n_1\rangle \to \langle m'\rangle$ and $F_2^{n_2}(s'_{2,1}, \ldots, s'_{2,n_2}) : \langle n_2\rangle \to \langle m'\rangle$ form a commuting square that is also an RPO, in fact any other pair of arrows, forming a commuting square factorizing the original one, needs to be in the form $F_1^{n_1}(s''_{1,1}, \ldots, s''_{1,n_1}) : \langle n_1\rangle \to \langle m''\rangle$ and $F_2^{n_2}(s''_{2,1}, \ldots, s''_{2,n_2}) : \langle n_1\rangle \to \langle m''\rangle$, with the two sequences $\langle s''_{1,i}\rangle$ and $\langle s''_{2,i}\rangle$ defining a unifier for $t_1, t_2$. The unique arrow factorizing the two commuting squares is

$F^{m'}(s'''_1, \dots s'''_{m'})$, with $\langle s'''_i \rangle$ given by the mgu property.




(iii) In this point we consider all the remaining cases, that is where: $s_1 = c_1(s_{1,1}, \dots, s_{1,k_1})$, $s_2 = c_2(s_{2,1}, \dots, s_{2,k_2})$ and either $F_1^{n_1}$ and $F_2^{n_2}$ do not appear in corresponding subterms, or $c_1 = F_1^{n_1}$ or $c_2 = F_2^{n_2}$. Let us consider the term $s'_1$ obtained from $s_1$ by substituting any maximal subterm $s_o$ not containing $F_1^{n_1}$ by a ground variable $X_{s_o}$.
For example, if $s_1 = c_1(s_{1,1}, c_2(s_{1,2,1}, F_1^{n_1}(s_{1,2,2,1}, s_{1,2,2,2}), s_{1,2,3}))$ then $s'_1$ is the term $c_1(X_{s_{1,1}}, c_2(X_{s_{1,2,1}}, F_1^{n_1}(X_{s_{1,2,2,1}}, X_{s_{1,2,2,2}}), X_{s_{1,2,3}}))$, and analogously for the term $s_2$. Let $s''_1 = s'_1 \circ t_1$, and $s''_1 = s'_2 \circ t_2$. Now we have that: $s''_1[s_{1,\vec{l_1}}/X_{s_{1,\vec{l_1}}}, \dots, s_{1,\vec{l_{m_1}}}/X_{s_{1,\vec{l_{m_1}}}}] = s''_2[s_{2,\vec{j_1}}/X_{s_{1,\vec{j_1}}}, \dots, s_{1,\vec{j_{m_2}}}/X_{s_{1,\vec{j_{m_2}}}}]$ that is, there exists a unifier, we can consider the most general unifier, given by a pair tuples of terms $s'_{1,\vec{l_1}}, \dots, s_{1,\vec{l_{m_1}}}$ and $s_{2,\vec{j_1}}, \dots, s_{1,\vec{j_{m_2}}}$. By repeating the arguments used at point (ii), we have that $s'_1[s'_{1,\vec{l_1}}/X_{s_{1,\vec{l_1}}}, \dots, s'_{1,\vec{l_{m_1}}}/X_{s_{1,\vec{l_{m_1}}}}]$ and $s'_2[s'_{2,\vec{j_1}}/X_{s_{1,\vec{j_1}}}, \dots, s'_{1,\vec{j_{m_2}}}/X_{s_{1,\vec{j_{m_2}}}}]$ form an RPO. □

6.1. **CL* as Second-order Rewriting System.** In this section, we consider the second-order context category for the combinatory calculus CL* and we show that the weak IPO bisimilarities thus obtained coincide with the observational equivalences on $\lambda$-calculus. Interestingly, the second-order open bisimilarity gives a uniform characterization also on open terms.

Note that the terms of CL are defined by the signature $\Sigma_{CL} = \{K, S, \mathtt{app}\}$, where `app` is the binary operation of application that is usually omitted. So the term **SKK** actually stands for $\mathtt{app}(\mathtt{app}(\mathbf{S}, \mathbf{K}), \mathbf{K})$.

First we deal with the lazy case, then we will sketch also the cbv case.

6.1.1. *The Lazy Second-order Reactive System.*

**Definition 6.3** (Lazy Second-order Reactive System on CL*)**.** The *lazy second-order reactive system* $\mathbf{C}_l^{2*}$ consists of:

- the category whose objects are the lists with at most one element, and whose arrows $\epsilon \to \langle n \rangle$ are the terms of CL* with, at most, $n$ (first order) metavariables,

$$M^n \ ::= \ X_1 \mid \dots \mid X_n \mid \mathbf{K} \mid \mathbf{S} \mid \mathbf{K}' M^n \mid \mathbf{S}' M^n \mid \mathbf{S}'' M^n M^n \mid M^n M^n$$

  and whose arrows $\langle m \rangle \to \langle n \rangle$ are the second-order contexts defined by:

$$\mathbb{C}^{m,n} \ ::= \ F(M_1^n, \dots, M_m^n) \mid M^n \mathbb{C}^{m,n} \mid \mathbb{C}^{m,n} M^n$$

- the reactive contexts are all the second-order applicative contexts of the shape $F(M_1^n, \dots, M_m^n) N_1^n \dots N_k^n$;
- the reaction rules are
  $\mathbf{K}X_1 \to \mathbf{K}'X_1 \qquad \mathbf{K}'X_1X_2 \to X_1$
  $\mathbf{S}X_1 \to \mathbf{S}'X_1 \qquad \mathbf{S}'X_1X_2 \to \mathbf{S}''X_1X_2 \qquad \mathbf{S}''X_1X_2X_3 \to (X_1X_3)(X_2X_3)$
  where $\mathbf{K}X_1, \mathbf{S}X_1 : \epsilon \to \langle 1 \rangle$, $\mathbf{K}'X_1X_2, \mathbf{S}'X_1X_2 : \epsilon \to \langle 2 \rangle$ and $\mathbf{S}''X_1X_2X_3 : \epsilon \to \langle 3 \rangle$.

Second-order contexts as defined above can be represented by $C[F(M_1, \dots, M_m)]$, where $C[\,]$ is a unary first-order context on CL$^*$ (with metavariables). To maintain the notation for contexts used in Sections 4, 5, in the sequel a second-order context $C[F(M_1, \dots, M_m)] : \langle m \rangle \to \langle n \rangle$ will be more conveniently written as $C[\,]_\theta$, where $\theta$ is a substitution s.t. $\theta(X_i) = M_i$ for all $i = 1, \dots, m$, moreover we write $M \stackrel{C[\,]_\theta}{\to} M'$ iff $C[M\theta] \to M'$. Given Proposition 6.2, and the underlined RPOs construction, we have:

**Corollary 6.4.** *The reactive system $\mathbf{C}_l^{2*}$ has redex RPOs.*

*Example*: Let $M = XM_1$. Some of the IPO reductions of $M$ are the following:
$XM_1 \stackrel{[\,]_{\{\mathbf{K}/X\}}}{\longrightarrow} \mathbf{K}'M_1$; $\quad XM_1 \stackrel{[\,]_{\{\mathbf{K}'Y/X\}}}{\longrightarrow} Y$; $\quad XM_1 \stackrel{[\,]_{\{\mathbf{K}'/X\}}Y}{\longrightarrow} M_1$; $\quad XM_1 \stackrel{[\,]_{\{\mathbf{S}/X\}}}{\longrightarrow} \mathbf{S}'M_1$;
$XM_1 \stackrel{[\,]_{\{\mathbf{S}'Y/X\}}}{\longrightarrow} \mathbf{S}''YM_1$; $\quad XM_1 \stackrel{[\,]_{\{\mathbf{S}'/X\}}Y}{\longrightarrow} \mathbf{S}''M_1Y$; $\quad XM_1 \stackrel{[\,]_{\{\mathbf{S}''YZ/X\}}}{\longrightarrow} (YM_1)(ZM_1)$;
$XM_1 \stackrel{[\,]_{\{\mathbf{S}''Y/X\}}Z}{\longrightarrow} (YZ)(M_1Z)$; $\quad XM_1 \stackrel{[\,]_{\{\mathbf{S}''/X\}}YZ}{\longrightarrow} (M_1Z)(YZ)$; $\quad XM_1 \stackrel{[\,]_{\{\mathbf{K}_{Y/X}\}}}{\longrightarrow} \mathbf{K}'YM_1$;
$XM_1 \stackrel{[\,]_{\{\mathbf{K}_{Y_1Y_2/X}\}}}{\longrightarrow} \mathbf{K}'Y_1Y_2M_1$.
Notice that $[\,]_{\{\mathbf{K}_{Y_1 \dots Y_n/X}\}}$ is an IPO context for any $n$.

In general, the IPO contexts are summarized in Figure 5.

Using Proposition 3.5, we can prove that the weak IPO bisimilarity $\simeq_{lI}^{2*}$ is a congruence, and it has a simpler characterization in terms of applicative contexts. Namely, we can consider as list extension category the category of second-order linear term contexts. In the alternative notation, a second order linear term contexts can be written as $C[\_{\theta_1}, \dots, \_{\theta_n}]$, where $C[\_1, \dots, \_n]$ is a first-order multi-holed context and $\theta_1, \dots, \theta_n$ are n substitutions, each one acting on the term put in the corresponding hole. By repeating the arguments for the first-order case, one can show that any second-order linear term context either is IPO uniform or it has a reactive index. Then, by Proposition 3.5, we have:

**Proposition 6.5.**

(i) *For all terms of $CL^*$ $M, N$, for any substitution $\theta$ and for any (possibly open) first-order context $C[\,]$,*

$$M \simeq_{lI}^{2*} N \implies C[M\theta] \simeq_{lI}^{2*} C[N\theta] \; .$$

(ii) $\simeq_{lI}^{2*} = \simeq_{lR}^{2*}$, *where $\simeq_{lR}^{2*}$ denotes the weak IPO bisimilarity, where only reactive IPO contexts are considered (see Figure 5).*

By Proposition 6.5(ii) above, the notion of IPO bisimilarity turns out to be much simpler, but it is still infinitely branching (when the term is of the shape $XP_0\vec{P}$ we have infinitely many IPO contexts $[\,]_{\{\mathbf{A}\vec{Y}/X\}}$). However, one can prove that also the contexts $[\,]_{\{\mathbf{A}\vec{Y}/X\}}$, for any $|\vec{Y}| \geq 1$ can be eliminated. This requires an "ad-hoc" reasoning:

| term $M$ | IPO contexts | reactive IPO contexts |
|---|---|---|
| $X$ | $[\,]_{\{\mathbf{A}Y/X\}}$, $[\,]_{\{\mathbf{A}/X\}}Y$, $\mathbf{A}\vec{Y}C_1[\,]_\emptyset$ | $[\,]_{\{\mathbf{A}Y/X\}}$, $[\,]_{\{\mathbf{A}/X\}}Y$ |
| $XP_0\vec{P}$ | $[\,]_{\{\mathbf{A}\vec{Y}/X\}}$, $\mathbf{A}\vec{Y}C_1[\,]_\emptyset$ | $[\,]_{\{\mathbf{A}\vec{Y}/X\}}$ |
| $\mathbf{C}\vec{P}$ , $M$ value | $[\,]_\emptyset X$, $\mathbf{A}\vec{Y}C_1[\,]_\emptyset$ | $[\,]_\emptyset X$ |
| $\mathbf{C}\vec{P}$, $M$ not value | $[\,]_\emptyset$, $\mathbf{A}\vec{Y}C_1[\,]_\emptyset$ | $[\,]_\emptyset$ |

where

$$\mathbf{A} \in \{\mathbf{K}, \mathbf{S}, \mathbf{K}'Z_1, \mathbf{S}'Z_1, \mathbf{S}''Z_1Z_2 \mid Z_1, Z_2 \text{ fresh}\}$$
$$\mathbf{C} \in \{\mathbf{K}, \mathbf{S}, \mathbf{K}', \mathbf{S}', \mathbf{S}''\}$$
$$C_1[\,] \text{ ranges over } C[\,] \ ::= \ [\,] \mid C[\,]Z \mid ZC[\,]$$

Figure 5: Second-order IPO contexts for the lazy CL$^*$.

| term $M$ | IPO contexts |
|---|---|
| $X$ | $[\,]_{\{\mathbf{A}/X\}}Y$ |
| $XP_0\vec{P}$ | $[\,]_{\{\mathbf{A}/X\}}$ |
| $\mathbf{C}\vec{P}$ , $M$ value | $[\,]_\emptyset X$ |
| $\mathbf{C}\vec{P}$ , $M$ not value | $[\,]_\emptyset$ |

where

$$\mathbf{A} \in \{\mathbf{K}, \mathbf{S}, \mathbf{K}'Z_1, \mathbf{S}'Z_1, \mathbf{S}''Z_1Z_2 \mid Z_1, Z_2 \text{ fresh}\}$$
$$\mathbf{C} \in \{\mathbf{K}, \mathbf{S}, \mathbf{K}', \mathbf{S}', \mathbf{S}''\}$$

Figure 6: Finitely branching second-order IPO contexts for the lazy CL$^*$.

**Proposition 6.6.** *The lazy weak IPO bisimilarity $\simeq^{2*}_{lI}$ has a finitely branching characterization in terms of the second-order IPO contexts of Figure 6.*

*Proof.* (sketch) Let $\simeq^{2*}_{lF}$ be the reduced bisimilarity obtained from $\simeq^{2*}_{lR}$ by not considering the contexts $[\,]_{\{\mathbf{A}\vec{Y}/X\}}$, for any $|\vec{Y}| \geq 1$. Then $\simeq^{2*}_{lR} \subseteq \simeq^{2*}_{lF}$. In order to show the converse, one can first prove that the following is a weak IPO bisimulation: $R = \{(M', N') \mid \exists\theta.\ (M' \frown M\theta \wedge N' \frown N\theta \wedge M \simeq^{2*}_{lF} N\}$, where $M \frown N$ means that $M$ and $N$ are $KS$-convertible. □

Finally, we are left to prove that the second-order weak IPO bisimilarity exactly recover the lazy observational equivalence. More in general, we will prove that the two equivalences coincide on open terms, when small variables of CL$^*$ are viewed as capital metavariables on the second-order reactive system. This gives a uniform finitely branching characterization of the observational equivalence on all terms.

**Proposition 6.7.** *For all $M, N \in \Lambda$, $M \approx_l N \iff \mathcal{T}(M) \simeq^{2*}_{lI} \mathcal{T}(N)$.*

**Proof of Proposition 6.7.** We will show that $\simeq^{2*}_{lI}$ coincides with the natural extension to open terms of the first-order IPO bisimilarity $\simeq^{*}_{lI}$ of Section 5.2.

**Definition 6.8.** Let $\widehat{\simeq}^*_{lI}$ be the extension of $\simeq^*_{lI}$ to open terms of CL* defined by, for all $M, N$ CL*-terms such that $FV(M), FV(N) \subseteq \{X_1, \ldots, X_n\}$,

$$M \widehat{\simeq}^*_{lI} N \text{ iff } \forall \theta : \{X_1, \ldots, X_n\} \to (CL^*)^0.\ M\theta \simeq^*_{lI} N\theta\ .$$

**Lemma 6.9.** *By identifying arrows in $\epsilon \to \langle n \rangle$ with open terms, $\simeq^{2*}_{lR} \subseteq \widehat{\simeq}^*_{lR}$.*

*Proof.* We show that $\mathcal{R} = \{(M\theta, N\theta) \mid M \simeq^{2*}_{lR} N \ \wedge\ M\theta, N\theta \in (CL^*)^0\}$ is a first-order bisimulation. From $M \simeq^{2*}_{lR} N$, by Proposition 6.5, we have $M\theta \simeq^{2*}_{lR} N\theta$. Assume $M\theta \xrightarrow{[\,]}_I M'$, since $M\theta \simeq^{2*}_{lI} N\theta$, then $N\theta \overset{[\,]}{\Rightarrow}_I N'$, $M' \simeq^{2*}_{RI} N'$ and $(M', N') \in \mathcal{R}$. Now assume $M\theta \xrightarrow{[\,]P}_I M'$, then $M\theta \xrightarrow{[\,]X}_I M''$ with $M''[P/X] = M'$, since $M\theta \simeq^{2*}_{lI} N\theta$ the also $N\theta \xrightarrow{[\,]X}_I N''$with $M'' \simeq^{2*}_{lI} N''$. Thus $N\theta \xrightarrow{[\,]P}_I N'$ and $N''[P/X] = N'$ is closed. Thus $(M', N') \in \mathcal{R}$. □

**Lemma 6.10.** *Let $M \in CL^*$, $M \to_l M'$. Then $M \widehat{\simeq}^*_{lI} M'$.*

*Proof.* The proof follows from the fact that $\forall \theta.\ M\theta \to^*_l M'\theta$ and $\simeq^*_{lI}$ is closed under $\to_l$. □

**Lemma 6.11.** $\widehat{\simeq}^*_{lR} \subseteq \simeq^{2*}_{lR}$.

*Proof.* We show that $\mathcal{R} = \{(M, N) \mid M \widehat{\simeq}^*_{lR} N\}$ is a second-order bisimulation. If $M \xrightarrow{[\,]_\theta \vec{X}}_I M'$, then there are two cases.
(i) $M = \mathbf{C}\vec{M}$, for a combinator $\mathbf{C}$ on $CL^*$. Then $\theta = \emptyset$, and for any closing $\theta$ and closed $\vec{P}$ such that $|\vec{X}| = |\vec{P}|$, $M\theta \xrightarrow{\vec{P}}_I M''$ and $M'' = M'\theta[\vec{P}/X]$. Since $M\theta \simeq^*_{lR} N\theta$, then $N\theta \xrightarrow{\vec{P}}_I N''$ and $M'' \simeq^*_{lR} N''$. There are two subcases: either $\vec{X} = [\,]$ or $\vec{X} = X$. In the first subcase, we have $M \to_I M'$ (second-order) and $N \Rightarrow N$ (second-order), thus by Lemma 6.10 $M' \widehat{\simeq}^*_{lR} N$, and hence $(M', N) \in \mathcal{R}$. In the second subcase, i.e. $\vec{X} = X$, $M = \mathbf{C}$, then one can check that also $N$ must be a combinator $\mathcal{C}'$, thus $N \xrightarrow{[\,]_\emptyset X} N'$ and $N'' = N'\theta[P/X]$. Thus $M' \widehat{\simeq}^*_{lR} N'$, and hence $(M', N') \in \mathcal{R}$.
(ii) $M = X\vec{M}$. Since for any closing $\theta$, $M\theta \simeq^*_{lR} N\theta$, then also $N \xrightarrow{[\,]_\theta \vec{X}}_I N'$. Moreover, for any $\overline{\theta}$ closing $M\theta, N\theta$, for any $\vec{P}$ such that $|\vec{P}| = |\vec{X}|$, we have $M\theta\overline{\theta} \xrightarrow{\vec{P}}_I M''$, $N\theta\overline{\theta} \xrightarrow{\vec{P}}_I N''$, $M'' = M'\theta[\vec{P}/\vec{X}]$, $N'' = N'\theta[\vec{P}/\vec{X}]$. Thus for all $\overline{\theta}'$. $M'\overline{\theta}' \simeq^*_{lR} N\overline{\theta}'$, hence $(M', N') \in R$. □

6.1.2. *The Cbv Second-order Reactive System.* The main difference between the cbv and the lazy case is that the variables in the cbv case are meant to represent values, consequently cbv substitutions have to map variables into values.

First of all, the values on CL* are defined by:

$$V \ ::= \ X \mid \mathbf{K} \mid \mathbf{K}'V \mid \mathbf{S} \mid \mathbf{S}'V \mid \mathbf{S}''VV\ .$$

**Definition 6.12** (Cbv Second-order Reactive System on CL*)**.** The *cbv second-order reactive system* $\mathbf{C}^{2*}_v$ consists of:

- the category whose objects are the lists with at most one element, and whose arrows $\epsilon \to \langle n \rangle$ are the terms of CL* with, at most, $n$ (first order) metavariables, and whose arrows $\langle m \rangle \to \langle n \rangle$ are the second-order contexts defined, briefly, by:

  $$\mathbb{C} \ ::= \ F(V_1, \ldots, V_m) \mid M\mathbb{C} \mid \mathbb{C}M$$

  where the values $V_1, \ldots, V_m$ and the term $N$ are built using $n$ variables.

| term $M$ | IPO contexts | reactive IPO contexts |
|---|---|---|
| $X$ | $[\,]_{\{\mathbf{A}/X\}}Y$, $\mathbf{A}[\,]_\emptyset$, $RC_1[\,]_\emptyset$, | $[\,]_{\{\mathbf{A}/X\}}Y$, $\mathbf{A}[\,]_\emptyset$ |
| $M$ a value but not a variable | $[\,]_\emptyset X$, $\mathbf{A}[\,]_\emptyset$, $RC_1[\,]_\emptyset$ | $[\,]_\emptyset X$, $\mathbf{A}[\,]_\emptyset$ |
| $M$ reducible | $[\,]_\emptyset$, $RC_1[\,]_\emptyset$ | $[\,]_\emptyset$ |
| $M$ contains a critical variable | $[\,]_{\{\mathbf{A}/Cr(M)\}}$, $RC_1[\,]_\emptyset$ | $[\,]_{\{\mathbf{A}/Cr(M)\}}$ |

where
$\mathbf{A} \in \{\mathbf{K}, \mathbf{S}, \mathbf{K}'X_1, \mathbf{S}'X_1, \mathbf{S}''X_1X_2 \mid X_1, X_2 \text{ fresh}\}$
$R$ ranges over $R \;::=\; \mathbf{A}Z \mid XR \mid RT$
$C_1[\,]$ ranges over $C[\,] \;::=\; [\,] \mid C[\,]T \mid TC[\,]$
with $T$ ranging over $T \;::=\; X \mid (TT)$

Figure 7: Second-order IPO contexts for cbv CL$^*$.

- the reactive contexts are defined by

$$\mathbb{D} \;::=\; F(V_1, \ldots, V_m) \mid \mathbb{D}M \mid V\mathbb{D}\;;$$

- the reaction rules are
  $\mathbf{K}X_1 \to \mathbf{K}'X_1 \qquad \mathbf{K}'X_1X_2 \to X_1$
  $\mathbf{S}X_1 \to \mathbf{S}'X_1 \qquad \mathbf{S}'X_1X_2 \to \mathbf{S}''X_1X_2 \qquad \mathbf{S}''X_1X_2X_3 \to (X_1X_2)(X_1X_3).$

By Proposition 6.2, we have:

**Corollary 6.13.** *The reactive system* $\mathbf{C}_v^{2*}$ *has redex RPOs.*

As in the lazy case, a second-order context $\mathbb{C} : \langle m\rangle \to \langle n\rangle$ will be more conveniently denoted by $C[\,]_\theta$, where $C[\,]$ is a unary first-order context and $\theta$ is a cbv substitution, i.e. s.t. $\theta(X_i)$ is a value, for all $i = 1, \ldots, m$.

According to our definition, there are terms that are neither values nor they are reducible (they do not contain any redex), the term $XY$ is an example. A term $M$ of this kind can be transformed in a reducible one by substituting a single specific variable with a value. We call *critical variable* a variable of this kind.

**Definition 6.14.** The *critical variable* of a second order term $M$, $Cr(M)$, if it exists, is recursively defined by:
$Cr(V) = \emptyset$,
$Cr(XV) = X$,
$Cr(VM) = Cr(M)$, if $V$ is a value different from a variable,
$Cr(MN) = Cr(M)$, if $M$ is not a value.

The second-order IPO contexts for cbv are summarized in Figure 7. In that figure, the symbol $R$ ranges over most general reducible terms. That is, any reducible term can be obtained by instantiating the variables of a term contained in that grammar. The symbol $T$ is used to represent general terms; remember that variables represent general values.

As for the previous case, by Proposition 3.5 and by considering the list extension category of by-value multi-holed second-order contexts we have:

**Proposition 6.15.**

| term $M$ | minimal IPO contexts |
|---|---|
| $X$ | $[\,]_{\{\mathbf{A}/X\}}Y$ |
| $M$ a value but not a variable | $[\,]_{\emptyset}X$ |
| $M$ reducible | $[\,]_{\emptyset}$ |
| $M$ contains a critical variable | $[\,]_{\{\mathbf{A}/Cr(M)\}}$ |

where
$\mathbf{A} \in \{\mathbf{K}, \mathbf{S}, \mathbf{K}'X_1, \mathbf{S}'X_1, \mathbf{S}''X_1X_2\}$

Figure 8: Minimal Second-order IPO contexts for cbv on CL*.

(i) *For all terms of $CL^*$ $M, N$, for any substitution $\theta$ and for any (possibly open) first-order context $C[\,]$,*

$$M \simeq^{2*}_{vI} N \implies C[M\theta] \simeq^{2*}_{vI} C[N\theta] \ .$$

(ii) $\simeq^{2*}_{vI} = \simeq^{2*}_{vR}$*, where* $\simeq^{2*}_{vR}$ *denotes the weak IPO bisimilarity, where only reactive IPO contexts are considered.*

It is important to notice that the reactive IPO contexts provide directly a finitely branching lts for the cbv combinatory logic (notice that, contrary to the lazy case, for the cbv case IPO contexts of the shape $[\,]_{\{\mathbf{A}\vec{Y}/X\}}$, for $|\vec{Y}| \geq 1$, do not exist, since substitutions have to map variables into values). However, using a direct argument, one should be able to further prune the lts by considering only the applicative contexts:

**Conjecture 6.16.** The cbv weak IPO bisimilarity $\simeq^{2*}_{vI}$ has a minimal characterization in terms of applicative reactive cbv second-order IPO contexts (see Figure 8).

The above conjecture can be seen as a Context Lemma for the second order IPO contexts. Presumably, it can be proved by adapting one of the techniques used to prove the Context Lemma for the cbv $\lambda$-calculus.

It is interesting to notice that Figure 8 can also be used to describe the finitely branching lts for lazy CL*. To do this explicitly, we only need to give the suitable, and obvious, definitions for "critical variable" and "value term" for the lazy case.

Finally, using an argument similar to the one used for the lazy case, one should be able to prove that the second-order weak cbv IPO bisimilarity exactly recovers the cbv observational equivalence, i.e.:

**Conjecture 6.17.** For all $M, N \in \Lambda$, $M \widetilde{\approx}_v N \iff \mathcal{T}(M) \simeq^{2*}_{vI} \mathcal{T}(N)$.

## 7. Final Remarks and Directions for Future Work

– There are several other attempts to deal with parametric rules in the literature. In particular, in [KSS05], the authors introduce the notion of *luxes* to generalize the RPO approach to cases where the rewriting rules are given by pairs of arrows having a domain different from 0. When instantiated to the category of contexts, the luxes approach allows to express rewriting rules not formed by pairs of ground terms but, instead formed by pairs of contexts (open terms), and so allowing parametricity. Compared to our approach, based on the notion of second-order context, the approach of luxes is more abstract and can be

applied to a wider range of cases (categories). However, if we compare the two approaches in the particular case of context categories, we find that luxes approach has a more restricted way to instantiate a given parametric rule. This restriction results in a not completely satisfactory treatment of the $\lambda$-calculus. It remains the open question of substituting the notion of second-order contexts with a more abstract and general one. This will allow to recover the extra generality of luxes.
– A possible alternative approach for dealing with the $\lambda$-calculus in Leifer-Milner's RPO setting, it that of using suitable encodings in the (bi)graph framework [Mil06]. However, we feel that our term solution based on second-order context categories and CL is simpler and more direct. Alternatively, in place of CL, one could also consider a $\lambda$-calculus with explicit substitutions, in order to obtain a convenient encoding of the $\beta$-rule, allowing for a representation as a second-order reactive system. This is an experiment to be done. Here we have chosen CL, since it is simpler; moreover, the correspondence between the standard $\lambda$-calculus and the one with explicit substitutions deserves further study.
– We have considered lazy and cbv strategies, however also other strategies, e.g. *head* and *normalizing* could be dealt with, possibly at the price of some complications due to the fact that such strategies are usually defined on open terms. It would be also interesting to explore non-deterministic strategies on $\lambda$-calculus.

## References


[AO93] S. Abramsky, L. Ong. Full Abstraction in the Lazy Lambda Calculus, *Information and Computation*, 105(2), 159–267, 1993.

[BKM06] F. Bonchi, B.Konig, U. Montanari. Saturated Semantics for Reactive Systems, *LICS'06*, IEEE, 69-80, 2006.

[BGK06] F. Bonchi, F. Gadducci, B. Konig. Process Bisimulation via a Graphical Encoding, *ICGT'06*, LNCS 4178, 168–183, 2006.

[CFC58] H. B. Curry, R. Feys, W. Craig. *Combinatory Logic*, Vol. 1. North–Holland, Amsterdam, 1958.

[DHL08] P. Di Gianantonio, F. Honsell, M. Lenisa. RPO, Second-order contexts, and $\lambda$-calculus. *FoSSaCS'08*, LNCS 398, 150–169, 2008.

[EHR92] L. Egidi, F. Honsell, S. Ronchi Della Rocca. Operational, Denotational and Logical Descriptions: a Case Study, *Fundamenta Informaticae*, 16(2), 149–169, 1992.

[EK06] H. Ehrig, B. König. Deriving bisimulation congruences in the DPO approach to graph rewriting with borrowed contexts. *Mathematical Structure in Computer Science*, 16(6), 1133–1163, 2006.

[GM05] F. Gadducci, U. Montanari. Observing Reductions in Nominal Calculi via a Graphical Encoding of Processes. *Processes, Terms and Cycles*, LNCS 3838, 106–126, 2005.

[HS86] R. Hindley, J. Seldin. Introduction to combinators and l-calculus, Cambridge University Press, 1986.

[KSS05] B. Klin, V. Sassone, P. Sobocinski. Labels from reductions: Towards a general theory. *CALCO'05*, LNCS 3629, 30–50, 2005.

[Lei01] J. Leifer. Operational congruences for reactive systems. *PhD thesis*, University of Cambridge Computer Laboratory, 2001.

[LM00] J. Leifer, R. Milner. Deriving bisimulation congruences for reactive systems. *CONCUR'00*, LNCS 1877, 243–258, 2000.

[Mil06] R. Milner. Local bigraphs and confluence: two conjectures. *Express'06*, ENTCS 175, 65–73, 2007.

[SS03] V. Sassone, P. Sobocinski. Deriving bisimulation congruences: 2-categories vs precategories. *FoSSaCS 2003*, LNCS 2620, 409–424, 2003.

[SS05] V. Sassone, P. Sobocinski. Reactive systems over cospans. *LICS'05*, IEEE, 311–320, 2005.

[Sew02] P. Sewell. From rewrite rules to bisimulation congruences. *Theoretical Computer Science*, 274(1-2), 183–230, 2002.

[Sob04] P. Sobocinski. Deriving process congruences from reduction rules. *PhD thesis*, University of Aarhus, 2004.