Distributive Evaluations of λ-calculus*

Cosimo LANEVE
Dip. Informatica, Universitá di Pisa, C.so Italia 40, Pisa, Italy †

Abstract

In this paper we address the problem of encoding evaluation strategies for the λ-calculus into prime event structures. In order for this to be possible the derivation spaces yielded by the evaluation mechanism must be prime algebraic cpo's. This requirement is not met by permutation equivalence (the standard concurrent semantics with which λ-calculus is equipped) since the derivation spaces it yields are upper semi-lattices. We solve this problem by taking the coarsest congruence contained in permutation equivalence such that permutations of disjoint reductions are equated and the downward closure of every derivation is a distributive lattice. This equivalence, called *distributive permutation equivalence*, is characterized directly by restricting permutations of redexes to those sets U which are distributive, i.e. for every $u \in U$, the development of every $V \subseteq (U \setminus \{u\})$ does not duplicate or delete u.

A simple consequence of our results is that the derivation spaces of the call-by-value λ-calculus are distributive lattices. Finally, we show that a sequential evaluation mechanism can not, in general, be effectively transformed into a maximally distributive one.

1 Introduction

Some recent papers (see [4, 8]) have given *concurrent semantics* to term rewriting systems, i.e. a semantics for the evaluation process which equate term rewritings differing just in the order of execution of independent reductions. The benefits of a concurrent semantics are clear as soon as one wants to implement the evaluation mechanism over a distributed system. By recognizing the independence relation, a concurrent semantics states the reductions which can be performed by different processors.

A concurrent semantics equates two derivations (i.e. sequences of reductions) if the second can be obtained from the first by exchanging the order of adjacent reductions which operate on "different" parts of the term. In concurrency theory this abstract notion of computation is neatly expressed through partial orders of reductions (or, in general, events). In particular, the order relation on the set of events describes causal dependency and its linearizations yield all the sequential computations which are equivalent to the original one. The sets of partial orders describe the states of the system (the amount of abstract computation already performed). Therefore states are naively provided with a lattice structure: the meet and join are respectively the intersection

*Partially supported by the ESPRIT Basic Research Project 6454 - CONFER.

†Current Address: INRIA-Sophia Antipolis (Project Meije), 2004 Route des Lucioles BP 93, 06902 Valbonne, France. E-mail: `cosimo@garfield.cma.fr`

and union of the events in the two partial orders. Furthermore, meets and joins hold the infinitely distributive property, which is inherited by the operations of intersection and union. In fact, distributive lattices reflect the *physical nature* of distributed systems. Two independent events e_1 and e_2 in such systems can be performed by two different processors which have no possibility of affecting each other. Therefore, if (as it is usual in concurrency theory) no assumption is made about the speeds of the two processors, the execution of the two events give rise to a lattice of configurations of the whole system corresponding to the states when the set of events that have taken place is $\emptyset$, $\{e_1\}$, $\{e_2\}$ or $\{e_1, e_2\}$.

Winskel shows that (algebraic) infinitely distributive lattices are exactly the domain of configurations of partial orders (or elementary event structures, in his terminology [10]). The generalization of this result to nondeterministic systems states that *prime event structures* (i.e. partial orders constrained by the relation of *consistency* plus the "finite causes" axiom, see [10]) give rise to sets of configurations which are (finitary) *prime algebraic domains*, that is algebraic complete partial orders such that the downward closure of every element is a distributive lattice. Prime event structures are a simple and general model to describe concurrent systems. The aim of this paper is to express the concurrent semantics of λ-calculus, a particular term rewriting system, by means of prime event structures. Indeed, it will be sufficient to turn the derivation spaces of λ-terms into prime algebraic cpo's, after which we can apply Winskel's results to obtain the required prime event structure. It should be straightforward to generalize our results to term rewriting systems; we talk about this generalization in the Conclusion.

There is a standard technique for giving a concurrent semantics to term rewriting systems which are left linear and non ambiguous (called *orthogonal* in [4]). The technique uses a well-known theorem, the *Parallel Move Theorem* (see [3] or Theorem 2.2 in Section 2), stating that all the development of a set of redexes in an expression t end at the same expression and have the same effect over redexes in t. Hence, the equivalence, named *permutation equivalence* is the congruence w.r.t. sequential composition of the axiom: "all the developments are equal".

The Parallel Move Theorem does no more hold as soon as the non ambiguity requirement is dropped. This is the case, for instance, of nondeterministic systems. The generalization of permutation equivalence to ambiguous term rewriting systems need some preliminary notion of *compatibility* among derivations (see [1]). Another approach overcoming this drawback is to define an axiomatization over derivations capturing permutation equivalence [8]. Actually, in [5], it is shown that the axiomatization which results from Meseguer's technique in general does not match with permutation equivalence due to the interplay of the rewriting rules with the axioms quotienting the term algebra. Therefore, case by case, a package of further axioms must be provided.

Anyway, more dramatic consequences follow by observing that partial orders are no longer suitable to modelling computations which are permutation equivalent. This is caused by the following fact:

a reduction can be cloned or can disappear due to the execution of a parallel reduction.

As a result of this phenomenon, permutation equivalence equates derivations yielding a different number of reductions. Two examples in λ-calculus will clarify the point. Let $\mathbf{I} = \lambda x.\, x$ and consider the term $t = (\lambda x.\, xx)(\mathbf{I}z)$. The term t has two redexes to which two different reductions correspond. The reductions can be uniquely identified by the access paths in the syntactic tree of the corresponding redexes (see Section 2). The access path of the reduction corresponding to the outermost redex is ε whilst that corresponding to the redex $\mathbf{I}z$ is 2. Now, if the redex 2 is fired before the redex ε there is no side effect over ε, meaning that the redex ε must still be fired. On the other hand, if the redex ε is fired before the redex 2, the redex 2 is duplicated due to the multiple occurrence of the variable x inside the body of the abstraction $\lambda x.\, xx$ and the derivation $\sigma = t \xrightarrow{2} (\lambda x.\, xx)z \xrightarrow{\varepsilon} zz$ turns out to be equivalent to $\rho = t \xrightarrow{\varepsilon} (\mathbf{I}z)(\mathbf{I}z) \xrightarrow{1} z(\mathbf{I}z) \xrightarrow{2} zz$ (notation $\sigma \equiv \rho$) and to $t \xrightarrow{\varepsilon} (\mathbf{I}z)(\mathbf{I}z) \xrightarrow{2} (\mathbf{I}z)z \xrightarrow{1} zz$. A similar situation results by examining the derivations of the term $t' = (\lambda x.\, ((\lambda y.\, z)(xw)))\mathbf{I}$. Due to the disappearance of redexes, the two derivations $t' \xrightarrow{12} (\lambda x.\, z)\mathbf{I} \xrightarrow{\varepsilon} z$ and $t' \xrightarrow{\varepsilon} (\lambda y.\, z)(\mathbf{I}w) \xrightarrow{2} (\lambda y.\, z)w \xrightarrow{\varepsilon} z$ are identified.

More insights can be obtained by considering the *derivation spaces* of λ-terms. These are the set of coinitial computations ordered by the *embedding relation*, informally defined as "a derivation σ is *embedded* into ρ if there exists a third derivation τ such that $\sigma; \tau \equiv \rho$". The derivation spaces of the terms t and t' are illustrated in Figure 1. Notice that, even though in this simple case the partial order is a lattice, it is not

$a: \quad (\lambda x.\, xx)(\mathbf{I}z)$ $\qquad\qquad$ $b: \quad (\lambda x.\, ((\lambda y.\, z)(xw)))\mathbf{I}$

Figure 1: Derivation space of two λ-terms.

distributive (for instance, in Figure 1.a, $\varepsilon \sqcup (\varepsilon; 2 \sqcap 2) \neq (\varepsilon \sqcup \varepsilon; 2) \sqcap (\varepsilon \sqcup 2)$, where we have forgotten terms in reductions).

The above observation makes us definitely reject permutation equivalence. In this paper, we restrict the term rewriting system to be the λ-calculus and we take the coarsest congruence contained in permutation equivalence which equates permutations of disjoint reductions and is such that the downward closure of every derivation is a distributive lattice. The constraint that a concurrent semantics must allow *at least* permutations of disjoint reductions seems to be a minimal requirement due to the fact that disjoint reductions change disjoint parts of the term and can not possibly affect each other.

The equivalence in question, called *distributive permutation equivalence*, is characterized directly in Section 3 by restricting permutations of redexes to those sets U which are distributive, i.e. for every $u \in U$, the development of every $V \subseteq (U/\{u\})$ does not duplicate or delete u.

In Section 4, as well as proving that the two characterizations of distributive permutation equivalence match, we show that derivation spaces yielded by the call-by-value λ-calculus are distributive lattices. This is achieved simply by recognizing that the set of redexes offered, at every term, by the call-by-value λ-calculus is always distributive.

Section 5 is devoted to the problem of finding the maximal distributivity for a (sequential) evaluation mechanism, that is, given a reduction strategy $\mathcal{N}$ and a term t, find the maximal set of distributive redexes at t which are fired by $\mathcal{N}$. We prove that, in general, turning a reduction strategy into a maximally distributive one is not recursive.

2 Permutation equivalence

We recall here, mainly from [7], the preliminary notions for the remaining part of the paper.

Let $\mathcal{V}$ be a denumerable set of *variables*, ranged over by x, y, z, ... The set of λ-terms over $\mathcal{V}$, ranged over by $t, t_1, \cdots$, is the least set containing $\mathcal{V}$ and closed under the operations of *abstraction* $\lambda x.\, t$ and *application* $t_1 t_2$.

We will use round brackets for making explicit the order of applications. For example $(t_1 t_2) t_3$ means that t_3 applies to the result of the application of t_1 to t_2.

As usual, λ-terms, are quotiented by the least congruence generated by the axiom α below

$$\lambda x.\, t \;=\; \lambda y.\, t\{^y/_x\} \qquad \textit{y does not occur free in t}$$

where the operation of *substitution* $t\{^y/_x\}$ and the notion of *free variable* are taken to be the standard ones. The unique rewriting rule is β-reduction:

$$(\lambda x.\, t)t' \;\rightarrow\; t\{^{t'}/_x\}$$

Subexpressions of terms are denoted by their access paths in the syntactic tree representing the term. Access paths are ranged over by u, v, w, $\cdots$ and are partially ordered by the prefix order: $u \preceq w$ iff there exists v such that $uv = w$. If $u \not\preceq w$ and $w \not\preceq u$ then we call u and w *disjoint* and write $u|w$. Of course, we have $u \prec w$ if $u \preceq w$ and $u \neq w$.

Let us call *redex* a term of the form $(\lambda x.\, t)t'$. The set $\mathcal{O}_\beta(t)$ will denote the set of redex occurrences in t while Throughout the paper we will assume that terms are taken up to α-conversion.

We say that t *reduces* to t' at occurrence u if t' is obtained from t by replacing the redex $(\lambda x.\, t_1)t_2$ at occurrence u with the instance of the right hand side of the β-conversion, namely $t_1\{^{t_2}/_x\}$. We use $\mathbf{u}$ to denote the *reduction* from t given by β-conversion of redex at u. Whenever it will be not clear from the context we will explicitly give the initial and ending terms of $\mathbf{u}$ by using the notation $t \xrightarrow{u} t'$. A *derivation* is a sequence $\mathbf{u}_1 \,;\, \mathbf{u}_2 \,;\, \ldots \,;\, \mathbf{u}_n$ of reductions such that the ending term of $\mathbf{u}_i$ is equal to the initial term of $\mathbf{u}_{i+1}$. Derivations will be ranged over by σ, ρ and the *empty reduction* from t to t is denoted by $\underline{\varepsilon}_t$ (we will omit the index t when it is clear from the context). Of course, as usual with strings, $\underline{\varepsilon}; \sigma = \sigma = \sigma; \underline{\varepsilon}$.

Let u a redex in t and σ a derivation starting from t. The redex u can be fired or duplicated or deleted during σ. The formal notion expressing "what happens at u after σ" is the residual mapping below. To this aim, firstly, let us define the set $o_x(t)$ of occurrences of the free variable x inside a term t. The definition goes by structural induction over t:

1. $o_x(y) \;=\; \begin{cases} \{\varepsilon\} & \textit{if } x = y \\ \varnothing & \textit{otherwise} \end{cases}$

2. $o_x(\lambda y.\, t) \;=\; \begin{cases} \varnothing & \textit{if } x = y \\ \{2u \mid u \in o_x(t)\} & \textit{otherwise} \end{cases}$

3. $o_x(t_1 t_2) \ = \ \{1u \mid u \in o_x(t_1)\} \cup \{2u \mid u \in o_x(t_2)\}$.

Definition 2.1 Let $t \xrightarrow{u} t'$, $v \in \mathcal{O}_\beta(t)$ and $(\lambda x.\, t_1)t_2$ be the subexpression of t at u. The set $v/\mathbf{u}$ of *residuals* of v by $\mathbf{u}$ is the following subset of $\mathcal{O}_\beta(t')$:

$$v/\mathbf{u} \ = \ \begin{cases} \varnothing & \textit{if } u = v \\ \{v\} & \textit{if } \ u|v \ \mathbf{or} \ v \prec u \\ \{uw\} & \textit{if } v = u12w \\ \{uwv_1 \mid w \in o_x(t_1)\} & \textit{if } v = u2v_1 \end{cases}$$

■

The generalization of the above definition to derivations is given inductively by:

1. $u/\underline{\varepsilon} \ = \ \{u\}$;

2. $u/(\sigma\,;\mathbf{v}) \ = \ \{w \mid w \in v'/\mathbf{v} \ \ \wedge \ \ v' \in u/\sigma\}$.

We remark that, by definition, $u/(\sigma\,;\mathbf{v}) \ = \ (u/\sigma)/\mathbf{v}$ and, in general, $u/(\sigma\,;\sigma') \ = \ (u/\sigma)/\sigma'$.

Let t be a λ-term, $U \subseteq \mathcal{O}_\beta(t)$ and $\sigma \ = \ \mathbf{u}_1\,;\mathbf{u}_2\,;\ldots\,;\mathbf{u}_n\,;\cdots$ be a derivation starting at t. The σ is *relative* to U iff for every $n \geq 1$, $u_n \in U/(\mathbf{u}_1\,;\cdots\,;\mathbf{u}_{n-1})$. The sequence σ is a *development* of U iff σ is relative to U and $U/\sigma \ = \ \varnothing$. In order to generalize residual calculi to pairs of derivations, we need another kind of reductions: *parallel reductions*. To this aim we are supported by the theorem below.

Theorem 2.2 (**The Parallel Move Theorem** [3]) Let $U \subseteq \mathcal{O}_\beta(t)$ for some λ-term t, then

1. there is no infinite derivation relative to U;

2. all developments end at the same expression;

3. for all $V \subseteq \mathcal{O}_\beta(t)$, if σ_1 and σ_2 are two developments of U, then $V/\sigma_1 \ = \ V/\sigma_2$. ■

Let us denote with $\mathbf{U}$ the parallel reduction firing at the same time the set U of redexes. The previous definition of derivation naively generalizes to derivations of parallel reductions. With a bit of inaccuracy we will write $\mathbf{u}$ for $\mathbf{U}$ whenever $U = \{u\}$. In this way derivations are a simple case of parallel derivations. We will denote with $\underline{\varnothing}_t$ the derivation firing an empty set of redexes (again, we will not hesitate to omit the index t as much as possible).

The residual of a set V after a (parallel) reduction $\mathbf{U}$, notation $V/\mathbf{U}$, is defined as $\bigcup_{v \in V} v/\sigma$, where σ is a development of U. Similar generalization concerns the residual of a set V after a (parallel) derivation σ. Theorem 2.2 can be generalized over parallel reductions. Namely, if U, V and W are sets of redexes in a term, then $\mathbf{U};\mathbf{V}/\mathbf{U}$ and $\mathbf{V};\mathbf{U}/\mathbf{V}$ terminates at the same expression and $W/(\mathbf{U};\mathbf{V}/\mathbf{U}) \ = \ W/(\mathbf{V};\mathbf{U}/\mathbf{V})$ (this last equality is known as the *Cube-Lemma*). Finally, the residual mapping can be lifted to derivations.

Notation Let σ be a derivations, we denote with $|\sigma|$ the number of (parallel) reductions in σ.

Let σ and ρ be coinitial derivations (i.e. they begin at the same term), the residual σ/ρ is a derivation starting at the end of ρ and defined inductively over $|\sigma| + |\rho|$ by:

$$\sigma/\rho \;=\; \begin{cases} \underline{\varepsilon} & \textit{if } \sigma = \underline{\varepsilon} \\ (\sigma'/\rho); U/(\rho/\sigma') & \textit{if } \sigma = \sigma'; \mathbf{U}. \end{cases}$$

Again, Theorem 2.2 can also be generalized over derivations. That is, if σ, ρ and τ are coinitial derivations (i.e. they begin at the same term), then $\sigma;\rho/\sigma$ and $\rho;\sigma/\rho$ terminates at the same expression and $W/(\sigma;\rho/\sigma) = W/(\rho;\sigma/\rho)$.

Now we can define the operational model of the λ-calculus. This model describes the derivations which can be performed by evaluating λ-terms. The properties of *associativity* ($(\sigma\,;\,\rho)\,;\,\phi \;=\; \sigma\,;\,(\rho\,;\,\phi)$) and the *identity* ($\underline{\varepsilon};\sigma = \sigma = \sigma;\underline{\varepsilon}$) hold naively over derivations. This observation leads to the idea of modelling the derivations as arrows of a *category*.

Definition 2.3 The operational model $\mathcal{L}$ of the λ-calculus is the category where the set of *objects* are terms up to axiom α and the *arrows* are the derivations. Arrows are quotiented by the axiom

$$\underline{\varepsilon} \;=\; \underline{\varnothing}$$

■

This model is too intensional to equip λ-calculus with a concurrent semantics. For example, if $t_1 \xrightarrow{u} t_1'$ and $t_2 \xrightarrow{v} t_2'$ then the two derivations $t_1t_2 \xrightarrow{1u} t_1't_2 \xrightarrow{2v} t_1't_2'$ and $t_1t_2 \xrightarrow{2v} t_1t_2' \xrightarrow{1u} t_1't_2'$ are not equated by the above axiomatization even if one can be obtained from the other by exchanging the two independent steps involved. Therefore further equalities must be provided. Here is the standard notion of *concurrent derivation* for λ-terms.

Definition 2.4 The category $\mathcal{L}^{\equiv}$ is the category $\mathcal{L}$ where morphisms are quotiented by the following axiom: for every λ-term t and every $U_1 \cup U_2 \subseteq \mathcal{O}_\beta(t)$:

$$\mathbf{U}_1\,;\,\mathbf{U}_2/\mathbf{U}_1 \;=\; \mathbf{U}_2\,;\mathbf{U}_1/\mathbf{U}_2$$

The equivalence relation quotienting the arrows of $\mathcal{L}$ is called *permutation equivalence* (notation $\equiv$). ■

So, for instance, if $U = \{u_1, u_2, u_3\}$ be a set of disjoint redexes and $U_1 = \{u_1, u_2\}$ then $\mathbf{U};\mathbf{U}_1/\mathbf{U} = \mathbf{U};\underline{\varnothing}$ and $\mathbf{U}_1;\mathbf{U}/\mathbf{U}_1 = \mathbf{U}_1;\mathbf{u}_3$ are permutation equivalent. Notice also that they are also equivalent to $\mathbf{u}_3;\mathbf{U}_1$.

The (generalization to derivations of the) Cube-Lemma is the basis to show that permutation equivalence is a congruence w.r.t. the operation of residual, i.e. if $\sigma \equiv \sigma'$ and $\rho \equiv \rho'$ then $\sigma/\rho \equiv \sigma'/\rho'$. There is still another useful lemma.

Lemma 2.5 Let σ and ρ be coinitial derivations with $|\sigma| = m$ and $|\rho| = n$. Then

1. $\sigma \equiv \rho$ if and only if $\sigma/\rho = \underline{\varnothing}^m$ and $\rho/\sigma = \underline{\varnothing}^n$; ($\underline{\varnothing}^m = \underbrace{\underline{\varnothing};\cdots;\underline{\varnothing}}_{m\ times}$).
2. $\sigma \sqsubseteq \rho$ if and only if $\sigma/\rho = \underline{\varnothing}^m$. ■

Furthermore, permutation equivalence is a congruence w.r.t. the operations of abstractions and application. Indeed, let a λ-context $C[\;]$ be a term generated by the following syntax:

$$[\,] \quad|\quad \lambda x.\,C[\,] \quad|\quad t\,C[\,] \quad|\quad C[\,]\,t$$

i.e. λ-contexts are λ-terms with one hole. We say that an equivalence $\equiv_\star$ over derivations is a *congruence* if, whenever $\mathbf{u}_1;\cdots;\mathbf{u}_m \;\equiv_\star\; \mathbf{v}_1;\cdots;\mathbf{v}_n$, for every context $C[\;]$ whose access path of the hole is w, it holds $\mathbf{u}_1';\cdots;\mathbf{u}_m' \;\equiv_\star\; \mathbf{v}_1';\cdots;\mathbf{v}_n'$, where $u_i' = wu_i$ and $v_i' = wv_i$.

Fact 2.6 Permutation equivalence is a congruence. ■

Finally, permutation equivalence induces, in turn, a preorder between coinitial computations, called the *embedding relation* (notation $\sqsubset$) and defined as:

$$\rho \sqsubset \sigma \quad \textit{if and only if} \quad \exists \tau.\, \rho; \tau \equiv \sigma.$$

The embedding relation can be turned into a partial order (denoted by $\sqsubseteq$) by quotienting derivations by $\equiv$.

Now, let t be a λ-term and let $\mathcal{L}^{\equiv}(t)$ be the comma category of objects under t, i.e. all the coinitial computations starting from t; finally consider the partial order $(\mathcal{L}^{\equiv}(t), \sqsubseteq)$. There is a well-known result due to Lévy.

Theorem 2.7 [6] Let t be a λ-term, $(\mathcal{L}^{\equiv}(t), \sqsubseteq)$ is an *upper semi-lattice* whose operation of least upper bound $\sigma \sqcup \rho$ between σ and ρ is defined as:

$$\sigma \sqcup \rho \stackrel{def}{=} \sigma; (\rho/\sigma)$$ ■

Notice that, in the above definition, $\sigma \sqcup \rho$ is also equal to $\rho; (\sigma/\rho)$. The above theorem can be rephrased in category theory by saying that $\mathcal{L}^{\equiv}(t)$ has pushouts. In general $(\mathcal{L}^{\equiv}(t), \sqsubseteq)$ is not a lattice: two computations may not have the greatest lower bound. Take, for instance, the term $t = (\lambda x.\, \mathbf{K}_a(x\mathbf{Y}_f))\mathbf{K}_b$, where $\mathbf{K}_a = \lambda x.\, a$, $\mathbf{K}_b = \lambda x.\, b$ and $\mathbf{Y}_f = (\lambda x.\, f(xx))(\lambda x.\, f(xx))$. Let σ be the reduction firing the redex 12 and let ρ be the reduction $t \xrightarrow{\varepsilon} \mathbf{K}_a(\mathbf{K}_b\mathbf{Y}_f) \xrightarrow{2} \mathbf{K}_a b$. Then the derivations σ and ρ lack of a greatest lower bound. Notice also that $(\mathcal{L}^{\equiv}(t), \sqsubseteq)$ is generally not topped. Consider, for example, $\Omega = (\lambda x.xx)(\lambda x.xx)$ then $(\mathcal{L}^{\equiv} \uparrow \Omega, \sqsubseteq)$ is a total ordering without top (derivations always have a finite number of reductions).

3 Distributive permutation equivalence

We have argued in the Introduction that permutation equivalence equates too many computations. Therefore the idea is to define a stronger equivalence (i.e. equating less computations) in order to get derivation spaces fitting with Winskel's requirements, i.e. the downward closure of every derivation is a distributive lattice.

In order to define this new equivalence we need the notion of distributive set of redexes. Let $U \in \mathcal{O}_\beta(t)$, U is called *distributive* if for every $u \in U$ and every $V \subseteq (U \setminus \{u\})$, $\#(u/\zeta_V) = 1$, where ζ_V is a development of V (by Theorem 2.2 we do not lose generality by restricting to consider just one development of V) and $\#(-)$ denotes the cardinality of its argument. That is the firing of redexes which are in parallel with u does not affect the individuality of u. We remark that set of redexes which are pairwise distributive are not distributive. Consider, for instance, the term $(\lambda x.\, (\lambda y.\, yy)x)(\mathbf{I}z)$; the three redexes ε, 12, 2 are pairwise distributive but their union does not yield a distributive set.

This section is devoted to show that every equivalence permuting reductions firing distributive sets of redexes yield derivation spaces which are prime algebraic. Distributive permutation equivalence rises as the congruence allowing the greatest number of permutations.

Definition 3.1 Let Ψ be a property over set of redexes and take the following axiom: if U_1 and U_2 are sets of redexes in a λ-term such that U is Ψ*-distributive*, i.e. $U_1 \cup U_2$ is distributive and $\Psi(U_1 \cup U_2)$, then

$$\mathbf{U}_1 ; \mathbf{U}_2/\mathbf{U}_1 = \mathbf{U}_2 ; \mathbf{U}_1/\mathbf{U}_2$$

We will denote with $\equiv_\Psi$ the equivalence yielded by the above axiom over arrows of $\mathcal{L}$ and with $\mathcal{L}^\Psi$ the resulting category (so $\equiv_\Psi$ is a congruence w.r.t. ";"). When $\Psi = true$ the equivalence will be called *distributive permutation equivalence* (notation $\equiv_d$). ■

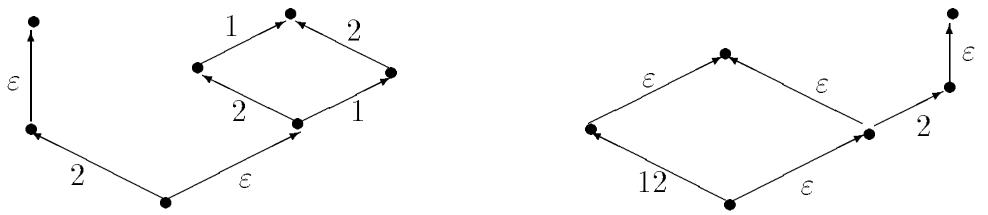
In the above definition Ψ also depends from the expression to which U_1 and U_2 belong. We have omitted such dependency in order to simplify the notation. We remark that, according to the above definition, distributive permutation equivalence is *finer* than permutation equivalence. For instance, take our running term $t = (\lambda x.\, xx)(\mathbf{I}z)$. The derivation $t \xrightarrow{2} (\lambda x.\, xx)z \xrightarrow{\varepsilon} zz$ is not $\equiv_d$-equivalent to $t \xrightarrow{\varepsilon} (\mathbf{I}z)(\mathbf{I}z) \xrightarrow{1} z(\mathbf{I}z) \xrightarrow{2} zz$ (whilst permutation equivalence identifies them). This example also shows that the equality $\rho; \sigma/\rho \equiv \sigma; \rho/\sigma$ no more holds for distributive permutation equivalence. As we will see in a moment it is true if and only if ρ and σ are *compatible* (i.e. there exists an upper bound of both).

Just as in the treatment of permutation equivalence, we can define the Ψ-*distributive embedding* relation (denoted $\sqsubset_\Psi$) between coinitial computations as

$$\rho \sqsubset_\Psi \sigma \quad \textit{if and only if} \quad \exists \tau.\, \rho; \tau \equiv_\Psi \sigma.$$

The Ψ-distributive embedding is a preorder with $\equiv_\Psi$ its associated equivalence. Let $\sqsubseteq_\Psi$ be the resulting partial order. The notation $\sqsubseteq_d$ will be reserved for the *distributive embedding* partial order.

Now, let t be a λ-term and let $\mathcal{L}^\Psi(t)$ be the comma category of objects under t. Firstly, we show the counterpart of Theorem 2.7 for $\sqsubseteq_\Psi$: the partial order $(\mathcal{L}^\Psi(t), \sqsubseteq_\Psi)$ is such that for every $\sigma \in \mathcal{L}^\Psi(t)$, its downward closure is an upper semi-lattice. Before embarking ourselves on the proofs, let us reconsider the λ-terms $t = (\lambda x.\, xx)(\mathbf{I}z)$ and $t' = (\lambda x.\, ((\lambda y.\, z)(xw)))\mathbf{I}$ in the Introduction. The differences between $\sqsubseteq$ and $\sqsubseteq_d$ are evident by comparing the derivation spaces of t and t' according to the embedding relation (as illustrated in Figure 1) and those according to the distributive embedding in Figure 2. Notice that the downward closure of every derivation is actually a distributive



a) $(\lambda x.\, xx)(\mathbf{I}z)$ *b)* $(\lambda x.\, ((\lambda y.\, z)(xw)))\mathbf{I}$

Figure 2: Derivation spaces according to $\sqsubseteq_d$.

lattice.

The fact that $(\mathcal{L}^\Psi(t), \sqsubseteq_\Psi)$ loses the property to be an upper semi-lattice must not scare too much the reader. Indeed the Church-Rosser property of λ-calculus is by no means lost: we can prove the validity of Church-Rosser property simply by providing a functor from $(\mathcal{L}^\Psi(t), \sqsubseteq_\Psi)$ to $(\mathcal{L}^\equiv(t), \sqsubseteq)$ which is the identity over objects and makes the identifications of permutation equivalence over arrows. That is, the more intensional nature of Ψ-distributive equivalence (w.r.t. permutation equivalence) concerns just derivations and not their i/o -behaviours.

Remark 3.2 Notice that, now, derivations are parallel derivations firing at each step a Ψ-distributive set of redexes. As c a consequence, the definition of σ/ρ will be no more total, namely it will be defined only when σ and ρ are compatible. This because, without the constraint of compatibility, when $\sigma = \sigma'; \mathbf{U}$, the set $U/(\rho/\sigma'$ is not, in general, Ψ-distributive.

Furthermore, in order to simplify matters in proofs, hereafter we will always consider derivations firing one redex at each step. When Ψ-distributive sets will be met, we will take a development of its. ■

The first result concerns a weaker form of the equality between $\sigma; \rho/\sigma$ and $\rho; \sigma/\rho$.

Proposition 3.3 Let σ, ρ and τ be coinitial derivations such that ρ, $\sigma \sqsubset_\Psi \tau$. Then σ/ρ and ρ/σ are well-defined, $\sigma; \rho/\sigma \equiv_\Psi \rho; \sigma/\rho$ and $\sigma; \rho/\sigma \sqsubset_\Psi \tau$.

Proof: By induction over $|\rho| + |\sigma|$.

(**basic case**) When $|\rho| + |\sigma| = 0$ the proposition holds trivially.

(**inductive case**) Assume that the proposition is true when $|\rho| + |\sigma| = n$; the case when the sum is $n + 1$ has two subcases, corresponding to $\sigma = \varepsilon$ and $\sigma = \sigma'; \mathbf{u}$. The first is immediate, for the latter let τ be the common upper bound of ρ and σ which exists by the assumption of compatibility. The inductive hypothesis guarantees that σ'/ρ and ρ/σ' are well-defined, $\sigma'; \rho/\sigma' \equiv_\Psi \rho; \sigma'/\rho$ and $\rho; \sigma'/\rho; \zeta \equiv_\Psi \tau$, for some ζ. The situation is illustrated in Figure 3.

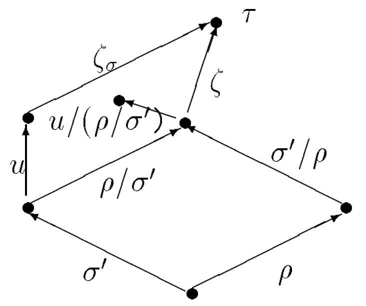


Figure 3: The situation of Proposition 3.3.

According to the definition of residual derivation and the inductive hypothesis, we have:

$$\rho; \sigma/\rho = \rho; \sigma'/\rho; u/(\rho/\sigma') \equiv_\Psi \sigma'; \rho/\sigma'; u/(\rho/\sigma')$$

Therefore, in order to prove that $\rho; \sigma/\rho \equiv_\Psi \sigma; \rho/\sigma$, it is enough to show that $\sigma'; \rho/\sigma'; u/(\rho/\sigma') \equiv_\Psi \sigma'; \mathbf{u}; \rho/\sigma$. To this aim we recall (see [7]) that $\rho/\sigma = (\rho/\sigma')/\mathbf{u}$, hence the proof reduces to that of $\rho/\sigma'; u/(\rho/\sigma') \equiv_\Psi \mathbf{u}; (\rho/\sigma')/\mathbf{u}$ (we have removed the prefix σ': this latter equation, under the hypothesis that $\equiv_\Psi$ is a congruence w.r.t. ';', gives $\sigma'; \rho/\sigma'; u/(\rho/\sigma') \equiv_\Psi \sigma'; \mathbf{u}; (\rho/\sigma')/\mathbf{u}$. See also Property 3.5 below).

There are two subcases: (**A**) $u/(\rho/\sigma') = \varnothing$ and (**B**) $u/(\rho/\sigma') \neq \varnothing$.

(**A**) In this case, the well-definition of σ/ρ is an obvious consequence of that of σ'/ρ. Thus let us concentrate on the proofs of $\sigma; \rho/\sigma \equiv_\Psi \rho; \sigma/\rho$ and $\sigma; \rho/\sigma \sqsubset_\Psi \tau$. We must show the existence of a derivation ξ such that $\rho/\sigma' \equiv_\Psi \mathbf{u}; \xi$. Let $\rho/\sigma' = \mathbf{u}_1; \cdots; \mathbf{u}_n$ and let k be the greatest natural under n such that, for every $1 \leq i \leq k-1$, $u/(\mathbf{u}_1; \cdots; \mathbf{u}_i)$ and u_{i+1} yield a distributive set. According to $\equiv_\Psi$, there is no way to permute the redexes $u/(\mathbf{u}_1; \cdots; \mathbf{u}_k)$ and u_{k+1} and, because $\sigma'; \mathbf{u}; \zeta_\sigma \equiv_\Psi \sigma'; \rho/\sigma'; \zeta$ and $u/(\rho/\sigma') = \varnothing$, the redex u must be fired before the $k+1$-th reduction. Hence let r, $1 \leq r \leq k-1$, be such that $u/(\mathbf{u}_1; \cdots; \mathbf{u}_{r-1}) = \{u_r\}$. By Ψ-distributive equivalence, we can permute the derivation $\mathbf{u}_1; \cdots; \mathbf{u}_r$ such that $\mathbf{u}_1; \cdots; \mathbf{u}_r \equiv_\Psi \mathbf{u}; \xi'$, for some ξ'. Observe that, by definition of residual derivation, we yield $\xi'; \mathbf{u}_{k+1}; \cdots; \mathbf{u}_n = (\rho/\sigma')/\mathbf{u}$.

By the above reasoning, it is immediate that both $\rho; \sigma/\rho$ and $\sigma; \rho/\sigma$ are (Ψ-distributively) embedded into τ.

(**B**) In order to prove that σ/ρ is well-defined we must check the Ψ-distributivity of $u/(\rho/\sigma')$. Let $\rho/\sigma' = \mathbf{u}_1; \cdots; \mathbf{u}_n$. Thus $u/(\rho/\sigma')$ is Ψ-distributive if for every i, $u/(\mathbf{u}_1; \cdots; \mathbf{u}_i)$ and u_{i+1} make a Ψ-distributive set. By contradiction, suppose that this is false and let k be the greatest natural such that $u/(\mathbf{u}_1; \cdots; \mathbf{u}_{k-1})$ and u_k yield a Ψ-distributive set. According to $\equiv_\Psi$, there is no way to permute the redexes $u/(\mathbf{u}_1; \cdots; \mathbf{u}_k)$ and u_{k+1} and, because $\sigma'; \mathbf{u}; \zeta_\sigma \equiv_\Psi \sigma'; \rho/\sigma'; \zeta$, the redex u must be fired before the $k+1$-th reduction. Moreover, due to the relation of Ψ-distributivity between $u/(\mathbf{u}_1; \cdots; \mathbf{u}_{j-1})$ and u_j $(1 \leq j \leq k)$, the firing of u before the $k+1$-th reduction causes $u/(\mathbf{u}_1; \cdots; \mathbf{u}_k) = \emptyset$. Thus, by definition of residual, $u/(\rho/\sigma') = \emptyset$. Contradiction. Therefore $u/(\rho/\sigma')$ is Ψ-distributive and σ/ρ is well-defined. The equality $\rho/\sigma'; \mathbf{u}/(\rho/\sigma') \equiv_\Psi \mathbf{u}; (\rho/\sigma')/\mathbf{u}$ is also a simple consequence of the proof.

It remains to show that it is possible to push ahead the reduction $\mathbf{u}/(\rho/\sigma')$ in the derivation ζ. Indeed, by arguing similarly to the derivation ρ/σ', we can prove that $\zeta \equiv_\Psi \mathbf{u}/(\rho/\sigma'); \zeta'$, for some ζ'. This latter observation yields $\sigma; \rho/\sigma; \zeta' \equiv_\Psi \tau$, which concludes the proof. ■

There is an immediate and useful consequence of the above proposition.

Corollary 3.4 Let ρ and σ be coinitial and compatible according to $\sqsubseteq_\Psi$, then $\sigma \sqsubseteq \rho$ implies $\sigma \sqsubseteq_\Psi \rho$.

Proof: By Proposition 3.3, $\sigma; \rho/\sigma \equiv_\Psi \rho; \sigma/\rho$. Furthermore the hypothesis and Lemma 2.5 imply that $\sigma/\rho = \underline{\emptyset}^m$ $(m = |\sigma|)$. Hence $\sigma; \rho/\sigma \equiv_\Psi \rho$, which, in turn, implies $\sigma \sqsubseteq_\Psi \rho$. ■

Property 3.5 (1) The equivalence $\equiv_\Psi$ is a congruence w.r.t. the residual operation. (2) $\equiv_\Psi$ is simplifiable "to the left", i.e. $\tau; \sigma \equiv_\Psi \tau; \rho$ if and only if $\sigma \equiv_\Psi \rho$.

Proof: (1) We must show that $\sigma \equiv_\Psi \rho$ implies that $\sigma/\tau \equiv_\Psi \rho/\tau$ (observe that the other half: $\sigma \equiv_\Psi \rho$ implies that $\tau/\sigma \equiv_\Psi \tau/\rho$ follows by the Cube-Lemma). Remark that we are assuming σ, ρ and τ compatible in order to guarantee the well-definition of σ/τ and ρ/τ. Let us prove that $(\sigma/\tau)/(\rho/\tau) \equiv_\Psi \underline{\varepsilon}$.

$$\begin{array}{rcll} (\sigma/\tau)/(\rho/\tau) & = & (\rho/\sigma)/(\tau/\sigma) & \text{by the Cube Lemma} \\ & \equiv_\Psi & \underline{\varepsilon}/(\tau/\sigma) & \text{by Corollary 3.4 and definition of } \equiv_\Psi \\ & = & \underline{\varepsilon} & \end{array}$$

Similarly for $(\rho/\tau)/(\sigma/\tau)$. Therefore, by Corollary 3.4 and Lemma 2.5, we yield the thesis.

(2) If $\sigma \equiv_\Psi \rho$ then, obviously, $\tau; \sigma \equiv_\Psi \tau; \rho$ because $\equiv_\Psi$ is a congruence w.r.t. sequential composition. *Vice versa*, observe that:

$$\begin{array}{rcl} (\tau;\sigma)/(\tau;\rho) & = & ((\tau;\sigma)/\tau)/\rho \\ & = & ((\tau/\tau)/(\sigma/(\tau/\tau)))/\rho \\ & = & (\underline{\varepsilon}/(\sigma/\underline{\varepsilon}))/\rho \\ & = & \sigma/\rho \end{array}$$

Thus, since $\tau; \sigma \equiv_\Psi \tau; \rho$ implies that $(\tau;\sigma)/(\tau;\rho) = \underline{\varepsilon}$, then $\sigma/\rho = \underline{\varepsilon}$. Symmetrically for ρ/σ. Finally, the compatibility of $\tau; \sigma$ and $\tau; \rho$ conjoined with Corollary 3.4 allow to use Lemma 2.5 also for $\equiv_\Psi$. Thus the property. ■

Remark 3.6 The above properties ensure that we can safely take the partial order $\sqsubseteq_\Psi$ instead of $\sqsubset_\Psi$. For instance, if $\rho \equiv_\Psi \sigma$ and $\rho \sqsubset_\Psi \tau$ then Property 3.5 implies $\sigma; \tau/\rho \equiv_\Psi \tau$. ∎

Let $\tau \in \mathcal{L}^\Psi(t)$, we denote with $\tau \Downarrow^\Psi$ the downward closure of τ w.r.t. $\sqsubseteq_\Psi$, i.e.

$$\tau \Downarrow^\Psi \;=\; \{\zeta \in \mathcal{L}^\Psi(t) \mid \zeta \sqsubseteq_\Psi \tau\}.$$

Corollary 3.7 Let $\tau \in \mathcal{L}^\Psi(t)$, then $(\tau \Downarrow^\Psi, \sqsubseteq_\Psi)$ is an upper semi-lattice.

Proof: Under the above hypothesis, by Proposition 3.3 we have that, for every pair ρ, $\sigma \in \tau \Downarrow^\Psi$, $\rho; \sigma/\rho \equiv_\Psi \sigma; \rho/\sigma$ and $\rho; \sigma/\rho$, $\sigma; \rho/\sigma \in \tau \Downarrow^\Psi$. It remains to show that $\rho; \sigma/\rho$ is the least upper bound of ρ and σ. By contradiction, assume the existence of an upper bound ζ of ρ and σ such that $\rho; \sigma/\rho \not\sqsubseteq_\Psi \zeta$. By Corollary 3.4, the fact that $\rho; \sigma/\rho$ is the least upper bound of ρ and σ in $\mathcal{L}^\equiv$ implies that $\rho; \sigma/\rho \sqsubseteq_\Psi \zeta$. Contradiction. ∎

A property of $\equiv_\Psi$ which clearly distinguishes it from $\equiv$ is that, if two derivations are $\equiv_\Psi$-equivalent, then they fire the same number of redexes. This recall a similar property of configurations of event structures, that if a configuration is reached by two different computations, the computations fire the same number of events. We define the *length* of a derivation σ (notation $\#(\sigma)$) inductively as follows:

$$\#(\sigma) \;=\; \begin{cases} 0 & \textit{if } \sigma = \underline{\varepsilon} \;\textbf{ or }\; \sigma = \underline{\varnothing} \\ \#(\sigma') & \textit{if } \sigma = \sigma'; \underline{\varnothing} \\ \#(\sigma') + 1 & \textit{if } \sigma = \sigma'; \mathbf{u} \;\textbf{ and }\; u \neq \varnothing \end{cases}$$

Notice that $\#(\sigma)$ is different from $|\sigma|$, the number of elementary derivations in σ. For instance $\#(\underline{\varepsilon}) = \#(\underline{\varnothing})$ whilst $|\underline{\varepsilon}| = 0$ and $|\underline{\varnothing}| = 1$. Here is the property reported above.

Fact 3.8 Let σ, ρ be two coinitial derivation; $\sigma \sqsubseteq_\Psi \rho$ implies $\#(\sigma) \leq \#(\rho)$.

Proof: By definition, if $\sigma \sqsubseteq_\Psi \rho$ then there exists a derivation τ such that $\sigma; \tau \equiv_\Psi \rho$. This guarantees that $\ell_\sigma \leq \ell_{\sigma;\tau}$. By definition of $\equiv_\Psi$, it is sufficient to show that if $U_1 \cup U_2 \subseteq \mathcal{O}_\beta(t)$ is distributive then $\#(\mathbf{U}_1; \mathbf{U}_2/\mathbf{U}_1) = \#(\mathbf{U}_2; \mathbf{U}_1/\mathbf{U}_2)$. Indeed, by definition of development, $\mathbf{U}_1 = \mathbf{u}_1; \cdots \mathbf{u}_n$ and $\mathbf{U}_2 = \mathbf{u}'_1; \cdots \mathbf{u}'_m$, where

$$\forall i \geq 1.\; [u_i \in U_1/\mathbf{u}_1; \cdots \mathbf{u}_{i-1} \quad \textbf{and} \quad u'_i \in U/\mathbf{u}'_1; \cdots \mathbf{u}'_{i-1}].$$

By the Ψ-distributivity of U_1 and U_2, every u_i, u'_i is empty or belongs to U_1 and U_2, respectively. Hence, since $\#$ is insensitive to empty reductions, $\#(\mathbf{U}_1) = \#(U_1)$ and $\#(\mathbf{U}_2) = \#(U_2)$ (the function $\#$ applied to sets gives the cardinality).

Back to the axiom of Definition 3.1, we have that Ψ-distributivity of $U_1 \cup U_2$ implies $\#(\mathbf{U}_1/\mathbf{U}_2) = \#(U_1 \setminus U_2)$ and $\#(\mathbf{U}_2/\mathbf{U}_1) = \#(U_2/U_1)$. The thesis follows by cardinality arguments. ∎

Since, for every t, $(\mathcal{L}^\Psi(t), \sqsubseteq_\Psi)$ is a partial order, Proposition 3.9 below is enough to show that it is a lattice. Theorem 3.11 proves that the lattice is distributive.

Proposition 3.9 If ρ and σ are two compatible coinitial derivations then they have greatest lower bound (notation $\rho \sqcap \sigma$).

Proof: By induction over $|\sigma| + |\rho|$.

(**basic case**) When $|\sigma| + |\rho| = 0$ the proposition is immediate.

(**inductive case**) Let $|\sigma| + |\rho| = n + 1$ and assume that the proposition holds when the

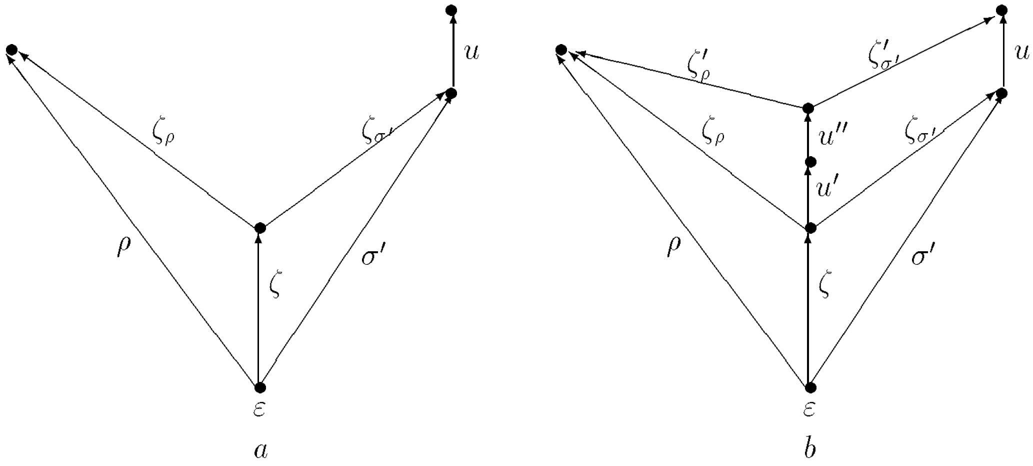


Figure 4: The situation of Proposition 3.9.

sum is n. Let $\sigma = \sigma'; \mathbf{u}$ and ζ be the greatest lower bound between σ' and ρ existing by inductive hypothesis. The case $u = \emptyset$ is trivial and, thus, omitted. Hence assume $u \neq \emptyset$. Let $\zeta; \zeta_{\sigma'} \equiv_\Psi \sigma'$ and $\zeta; \zeta_\rho \equiv_\Psi \rho$. The situation is illustrated in Figure 4.a. Let $\zeta_{\sigma'} = \mathbf{u}_1; \cdots; \mathbf{u}_n$. There are two cases:

(**A**) u can be (distributively) permuted with every u_i and $u'/\zeta_\rho = \emptyset$, where $u = u'/\zeta_{\sigma'}$;

(**B**) the dual of **A**.

(**A**) In this case, notice that $\zeta_{\sigma'}; \mathbf{u}$ results to be $\equiv_\Psi$-equivalent to $\mathbf{u}'; \zeta^+_{\sigma'}$ (for some $\zeta^+_{\sigma'}$). Due to the compatibility of $\zeta; \zeta_\rho$ and $\zeta; \zeta_{\sigma'}; \mathbf{u}$, by arguing similarly to Proposition 3.3, we can conclude that ζ_ρ is $\equiv_\Psi$-equivalent to $\mathbf{u}'; \zeta^+_\rho$ (for some ζ^+_ρ). We have to show that $\zeta; \mathbf{u}'$ is the greatest lower bound. Surely it is a lower bound and, for contradiction, assume that it is not the greatest. That is there exists $\zeta^\star$ such that $\zeta; \mathbf{u}'$ is strictly embedded into $\zeta^\star$. Notice that we can safely reduce to this case because, otherwise we can take $\zeta^\star \sqcup (\zeta; \mathbf{u}')$ which is still a lower bound of σ and ρ. Let us simplify the matters by getting $\zeta^\star = \zeta; \mathbf{u}'; \mathbf{u}''$ and let ζ'_ρ and $\zeta'_{\sigma'}$ be such that $\rho \equiv_\Psi \zeta^\star; \zeta'_\rho$ and $\sigma \equiv_\Psi \zeta^\star; \zeta'_{\sigma'}$. The situation is depicted in Figure 4.b.

According to the definitions, $\mathbf{u}'; \mathbf{u}''; \zeta'_{\sigma'} \equiv_\Psi \zeta_{\sigma'}; \mathbf{u}$, that is u'' can be (Ψ-distributively) permuted with u. Hence, let $\mathbf{u}^\star; \mathbf{u}^+ \equiv_\Psi \mathbf{u}'; \mathbf{u}''$ such that $\{u''\} = u^\star/\mathbf{u}'$ and $\{u^+\} = u'/\mathbf{u}^\star$. It follows that ζ is strictly embedded into $\zeta; \mathbf{u}^\star$ and, moreover, $\zeta; \mathbf{u}^\star$ is a lower bound of σ' and ρ. Contradiction.

(**B**) If it is not the case that u can be Ψ-distributively permuted with every u_i or, when this is possible, $u'/\zeta_\rho \neq \emptyset$ ($u = u'/\zeta_{\sigma'}$) then ζ is still the greatest lower bound of ρ and σ. Assume, for contradiction, that this is false. Again, let $\zeta; \mathbf{u}''$ be another lower bound of ρ and σ and let ζ'_ρ and $\zeta'_{\sigma'}$ be such that $\rho \equiv_\Psi \zeta; \mathbf{u}''; \zeta'_\rho$ and $\sigma \equiv_\Psi \zeta; \mathbf{u}''; \zeta'_\sigma$. As it stands, $u''/\zeta_\rho = u''/(\zeta_{\sigma'}; \mathbf{u}) = \emptyset$. Furthermore, by definition, $u''/(\zeta_{\sigma'}; \mathbf{u}) = \{w \mid w \in u_1/\mathbf{u} \ \wedge \ u_1 \in u''/\zeta_{\sigma'}\}$. Here there are two subcases:

B.1 u can not be Ψ-distributively permuted with the redexes fired by $\zeta_{\sigma'}$;

B.2 u can be Ψ-distributively permuted with the redexes fired by $\zeta_{\sigma'}$ and $u'/\zeta_\rho \neq \emptyset$.

(**B.1**) In this case, $\mathbf{u}''; \zeta'_\sigma \equiv_\Psi \zeta_{\sigma'}; \mathbf{u}$ implies that u'' must be fired inside $\zeta_{\sigma'}$, i.e. there exists a derivation ζ'' such that $\zeta_{\sigma'} \equiv \mathbf{u}''; \zeta''$, thus resulting $\zeta; \mathbf{u}''$ a lower bound of ρ and σ'. Contradiction.

(**B.2**) Here, observe that $u'/\zeta_\rho \neq \emptyset$ ($u = u'/\zeta_{\sigma'}$) whilst $u''/\zeta_\rho = \emptyset$. Therefore $u' \neq u''$. Moreover $\zeta_{\sigma'}; \mathbf{u} \equiv_\Psi \mathbf{u}''; \zeta'_\sigma$ implies that u'' must be fired inside $\zeta_{\sigma'}$. Hence a contradiction similar to the case **B.1** yields. ∎

The following property is needed to prove that the downward closures of $\sqsubseteq_\Psi$-derivation spaces are distributive.

Property 3.10 Let ρ, σ, τ be coinitial and compatible, then $\rho/\tau \sqcap \sigma/\tau \equiv_\Psi (\rho \sqcap \sigma)/\tau$.

Proof: It is straightforward to show that $(\rho \sqcap \sigma)/\tau \sqsubseteq_\Psi \rho/\tau \sqcap \sigma/\tau$. We will prove $\rho/\tau \sqcap \sigma/\tau \sqsubseteq_\Psi (\rho \sqcap \sigma)/\tau$ by contradiction. Assume the existence of a redex u in the final term of the derivation $(\rho \sqcap \sigma)/\tau$ such that $\rho/\tau \sqcap \sigma/\tau \equiv_\Psi (\rho \sqcap \sigma)/\tau; \mathbf{u}$ (the case of a sequence of redexes can be treated similarly). This situation is illustrated in Figure 5. According to the definitions:

$$\zeta_\rho; \tau/\rho \equiv_\Psi \tau/(\rho \sqcap \sigma); \mathbf{u}; \zeta'_\rho.$$

Notice that the redex u can not be fired during the derivation τ/ρ because, by definition of residual derivation, redexes fired by τ/ρ are exactly the residual redexes of τ after the derivation ρ (and this is not the case for u). Therefore, the redex u is fired during the derivation ζ_ρ. Of course, a similar reasoning holds for ζ_σ, too. Now, let $\tau/(\sigma \sqcap \rho) = \tau_{(1)}; \tau_{(2)}$ where $\tau_{(1)}$ is such that its final state t'_1 is the first λ-term in the derivation

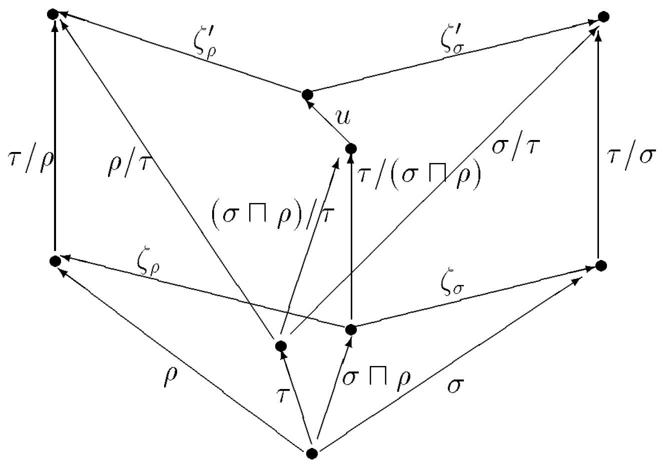


Figure 5: The situation of Property 3.10.

$\tau/(\sigma \sqcap \rho)$ with $u \in \mathcal{O}_\beta(t'_1)$ and such that u is distributive with every redex fired by $\tau_{(2)}$. We prove by induction over $|\tau_{(1)}|$ that $\sigma \sqcap \rho$ is not the greatest lower bound of ρ and σ, thus yielding a contradiction.

(**basic case**) When $|\tau_{(1)}| = 0$ we have that

$$\mathbf{u}'; (\tau/(\sigma \sqcap \rho))/\mathbf{u}'; \zeta'_\rho \equiv_\Psi \zeta_\rho; \tau/\rho.$$

where $u = u'/\tau_{(2)}$. Arguing similarly to Proposition 3.3, it is possible to yield the existence of a ζ^+_ρ such that $\zeta_\rho \equiv_\Psi \mathbf{u}'; \zeta^+_\rho$. Similarly for $\zeta_{\sigma'}$. Therefore $\sigma \sqcap \rho; \mathbf{u}$ is still a lower bound of σ and ρ, contrarily to the assumptions that $\sigma \sqcap \rho$ is the greatest one.

(**inductive case**) Assume that the thesis holds when $|\tau_{(1)}| = n$, let us show the case $|\tau_{(1)}| = n + 1$. We will prove that, if $\tau_{(1)} = \mathbf{u}_1; \ldots; \mathbf{u}_{n+1}$, then u_{n+1} is also a redex fired by ζ_ρ and ζ_σ and, therefore, we can apply inductive hypothesis to u_{n+1}. Let

$\tau'_{(1)} = \mathbf{u}_1; \ldots ; \mathbf{u}_n$, hence

$$\begin{array}{rcll} \zeta_\rho; \tau/\rho & \equiv_\Psi & \tau'_{(1)}; \mathbf{u}_{n+1}; \tau_{(2)}; \mathbf{u}; \zeta'_\rho & \\ & \equiv_\Psi & \tau'_{(1)}; \mathbf{u}_{n+1}; \mathbf{u}'; \tau'_{(2)}; \zeta'_\rho & \text{by hypothesis, for some } \tau'_{(2)}\ (u = u'/\tau_{(2)}) \end{array}$$

and, according to the definition of residual derivation (see [7]),

$$(\zeta_\rho/\tau'_{(1)}); (\tau/\rho)/(\tau'_{(1)}/\zeta_\rho) \ \equiv_\Psi \ \mathbf{u}_{n+1}; \mathbf{u}'; \tau'_{(2)}; \zeta'_\rho$$

It remains to show that the redex u_{n+1} must be fired inside $\zeta_\rho/\tau'_{(1)}$. Notice that, according to the hypothesis, u_{n+1} and u' can not be permuted. Thus, because u' must be fired by $\zeta_\rho/\tau'_{(1)}$ (for a reason similar to the basic case), it follows that u_{n+1} is fired by $\zeta_\rho/\tau'_{(1)}$, too. Q.e.d. ■

Theorem 3.11 Let $\zeta \in \mathcal{L}^\Psi(t)$, then $\zeta \Downarrow^\Psi$ is a distributive lattice.

Proof: Proposition 3.9 proves that $\zeta \Downarrow^\Psi$ is a lattice. As far as distributivity is concerned, let σ, ρ and τ be three coinitial derivations. We have to show the following equality:

$$\tau \sqcup (\sigma \sqcap \rho) \ \equiv_\Psi \ (\tau \sqcup \sigma) \sqcap (\tau \sqcup \rho)$$

Notice that its left hand side is equivalent to $\tau; (\sigma \sqcap \rho)/\tau$ and its right hand side is equivalent to $\tau; (\sigma/\tau \sqcap \rho/\tau)$. Hence it follows by Property 3.10 and by the fact that $\equiv_\Psi$ is preserved by ";". ■

Let us show an immediate implication of Theorem 3.11. In the Introduction we have argued that reduction strategies giving derivation spaces which are distributive lattices can be simulated by elementary event structures. The strategies having this property will be called *distributive*. Of course all one-step reduction strategies (lazy, leftmost-outermost, etc.), being deterministic, are distributive and can be faithfully encoded into a concurrent process. A well-known example of a reduction strategy showing up some parallelism is the *call-by-value* λ-calculus. Below we show that call-by-value λ-calculus is distributive. Actually it has been successfully encoded into π-calculus processes by Milner (see [9]): a formal comparison between Milner's encoding and the one raising from this paper goes outside the scope of this work.

Let $\mathcal{W}$, ranged over by v, be the set of *values* inductively defined by the rules:

$$x \quad | \quad \lambda x.\, t$$

where $x \in \mathcal{V}$ and t is a λ-term.

Definition 3.12 The *call-by-value* reduction relation $\rightarrow_v$ over λ-terms is the least relation satisfying the following rules:

$$\beta)\ \ (\lambda x.\, t)v \ \rightarrow_v \ t\{^v/_x\} \qquad ap_l)\ \ \frac{t_1 \ \rightarrow_v \ t'_1}{t_1 t_2 \ \rightarrow_v \ t'_1 t_2} \qquad ap_r)\ \ \frac{t_2 \ \rightarrow_v \ t'_2}{t_1 t_2 \ \rightarrow_v \ t_1 t'_2}$$ ■

A term $(\lambda x.\, t')v$ ($v \in \mathcal{W}$) is named a *call-by-value redex* of the λ-term t if, for some u, it is the subterm of t at occurrence u and there is no abstraction outside it. Let $\mathcal{O}^v_\beta(t)$ be the set of call-by-value redexes of a term t. Trivially, $\mathcal{O}^v_\beta(t) \subseteq \mathcal{O}_\beta(t)$ holds. Moreover observe that call-by-value λ-calculus is not a congruence w.r.t. abstraction and application: a call-by-value redex in t is no more a call-by-value redex in $\lambda x.\, t$.

Fact 3.13 Let $u, u' \in \mathcal{O}^v_\beta(t)$, with $u \neq u'$. Then $u|u'$.

Proof: According to the definition of $u|u'$ both $u \not\preceq u'$ and $u' \not\preceq u$ must hold. We show the first, the other follows with a symmetric reasoning. By contradiction, assume that $u \preceq u'$. Since, by hypothesis, $u \neq u'$, it remains the case $u \prec u'$. There are two possibilities: either $\exists u''.\, u' = u.1.u''$ or $\exists u''.\, u' = u.2.u''$. The first possibility can not hold because u' does not lies in an applicative context. The second is also vacuous because $u.2$ must be a value (it is an abstraction or a variable) and values can not show up call-by-value redexes. Contradiction. ■

Redexes that are disjoint are distributive, as stated by Fact 4.1 of the next Section. Therefore the following statement is a consequence of Theorem 3.11.

Corollary 3.14 The derivation spaces of the call-by-value λ-calculus are distributive lattices. ■

We end this section with two remarks. The first is that partial orders raising from $\sqsubseteq_d$ are not *coherent*, in general. Indeed, it is possible that a set of pairwise compatible derivations have no upper bound. As an example we recall the term given at the biginning of this section: $t = (\lambda x.\, (\lambda y.\, yy)x)(\mathbf{I}z)$. Notice that the three derivations firing the redexes ε, 12 and 2 are pairwise compatible but they have no upper bound. The second remark concerns completeness. Since derivations are finite, we lose the property that derivation spaces are complete partial order. As usual we can regain completeness by *ideal completion*: ideal completion of derivation spaces are finitary prime algebraic cpo's.

4 The coarsest equivalence theorem

This section is devoted to show a coarsest result concerning $\equiv_d$. To this aim, we consider the set of all equivalences contained into $\equiv$ and permuting *at least* redexes which are disjoint. Permutations of disjoint redexes seems to be a sort of *minimal* requirement for a concurrent semantics of the λ-calculus. This is due to the fact that disjoint reductions concern different parts of the term and can not affect each other. By the way, Fact 4.1 below shows that disjoint sets of redexes are distributive, too.

Formally, let $\equiv^\diamond$ be the equivalence over morphisms of the operational model $\mathcal{L}$ induced by the following axiom: for every term t and every pair $U_1,\, U_2 \subseteq \mathcal{O}_\beta(t)$ such that $U_1 \cup U_2$ is a set of mutually disjoint redexes,

$$\mathbf{U}_1\,;\,\mathbf{U}_2/\mathbf{U}_1 \;=\; \mathbf{U}_2\,;\,\mathbf{U}_1/\mathbf{U}_2.$$

Let $\sqsubseteq^\diamond$ be the partial order induced by $\equiv^\diamond$. We prove that, for every derivation τ, its downward closure w.r.t. $\sqsubseteq^\diamond$ is a distributive lattice. According to the definition of $\equiv^\diamond$, it is enough to show that every set of redexes which are mutually disjoint is distributive.

Fact 4.1 Let t be a λ-term and $U \in \mathcal{O}_\beta(t)$ be a set of mutually disjoint redexes. Then U is distributive.

Proof: Let $u \in U$ and $V \in (U/\{u\})$: we must show that if σ is a development of V then $\#(u/\sigma) = 1$. By definition of development, $\sigma = \mathbf{u}_1; \cdots; \mathbf{u}_n$ where, for every $1 \leq i \leq n$, $u_i \in V/\mathbf{u}_1; \cdots; \mathbf{u}_{i-1}$ and $V/\sigma = \emptyset$. Let $V_i = V/(\mathbf{u}; \cdots; \mathbf{u}_{i-1})$. Thus, the proof reduces to verify that, for every i, $u/(\mathbf{u}_1; \cdots; \mathbf{u}_{i-1}) = u$ and $\{u\} \cup V_i$ is a set of redexes which are mutually disjoint. By induction over i: the case $i = 0$ is immediate. The inductive case $i = k+1$ follows by noticing that $u'/u_{k+1} = u'$ when u' and u_{k+1} are disjoint. Therefore $(\{u\}\cup V_i)/(\mathbf{u}_1;\cdots;\mathbf{u}_{k+1}) = \{u\}\cup(V_i\setminus\{u_{k+1}\})$ which is a mutually disjoint set of redexes by inductive hypothesis. Q.e.d. ■

In the theorem which follows we will take a generic equivalence $\equiv_\star$ over morphisms in $\mathcal{L}$. Notice that $\equiv_\star$ must be a congruence w.r.t. sequential composition. Here is the coarsest equivalence theorem.

Theorem 4.2 Let $\equiv_\star$ be a congruence over morphisms of $\mathcal{L}$ and $\equiv^\diamond\subseteq\equiv_\star\subseteq\equiv$. If the partial order $\sqsubseteq_\star$ induced by $\equiv_\star$ holds the property that for every τ, $\tau\Downarrow^\star$ (the downward closure of τ w.r.t. $\sqsubseteq_\star$) is a distributive lattice, then $\equiv_\star\subseteq\equiv_d$.

Proof: By contradiction assume the theorem is false. Hence there exist two derivations ρ, σ such that $\rho\equiv_\star\sigma$ and $\rho\not\equiv_d\sigma$. Let $\rho\equiv_\star\rho_{(1)};\rho_{(2)}$ and $\sigma\equiv_\star\sigma_{(1)};\sigma_{(2)}$ such that $\ell_{\rho_{(1)}}$ is the largest among the derivations which are strictly embedded into ρ and σ satisfying the property that $\rho_{(1)}\equiv_d\sigma_{(1)}$ and $\rho_{(1)}\equiv_\star\sigma_{(1)}$.

Therefore, let $\rho_{(2)}=\mathbf{u};\rho'_{(2)}$ and $\sigma_{(2)}=\mathbf{u}';\sigma'_{(2)}$ (u, u' are not empty). There are two main cases, according to the set $\{u,\,u'\}$ being distributive or not.

(**A**) The set $\{u,\,u'\}$ is not distributive. Thus, by the hypothesis, $u\neq u'$, $u'/\rho_{(2)} = u/\sigma_{(2)}=\varnothing$. There are two basic cases (the others can be reduced to these):

(A.1) $|u/u'|>1$ and $|u'/u|=1$;

(A.2) $|u/u'|=0$ and $|u'/u|=1$.

(**A.1**) In this case, from $\rho\Downarrow^\star$ we can extract the sublattice in Figure 6.a. A well-known

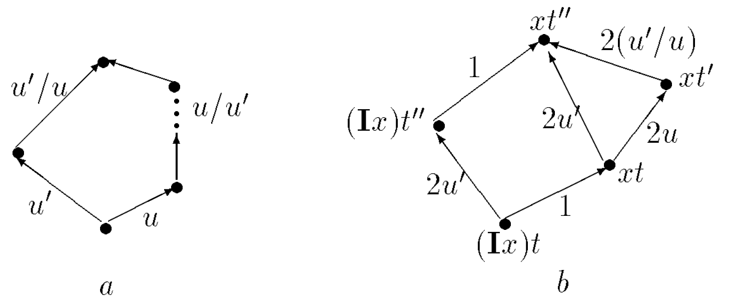


Figure 6: The situations of case **A** of Theorem 4.2.

result in lattice theory states that the presence in $\rho\Downarrow^\star$ of sublattices like the one in Figure 6.a makes $\rho\Downarrow^\star$ not distributive (see, for instance [2]).
(**A.2**) In this case, consider the term $(\mathbf{I}x)t$ (t is the initial term of the derivations $\rho_{(2)}$ and $\sigma_{(2)}$). By $\equiv^\diamond\subseteq\equiv_d$, the two derivations $\rho=(\mathbf{I}x)t\xrightarrow{1}xt\xrightarrow{2u'}xt''$ and $\sigma=(\mathbf{I}x)t\xrightarrow{2u'}(\mathbf{I}x)t''\xrightarrow{1}xt''$ must be $\equiv_\star$-equivalent. On the other hand, since $\equiv_\star$ is a congruence, we yield $xt\xrightarrow{2u'}xt''\equiv_\star xt\xrightarrow{2u}xt'\xrightarrow{2(u'/u)}xt''$. The situation is illustrated in Figure 6.b. Again, by arguing similarly to the subcase **A.1**, we deduce that $\rho\Downarrow^\star$ is not a distributive lattice.

(**B**) The set $\{u,\,u'\}$ is distributive. We remark that $\mathbf{u};\mathbf{u}^+\not\equiv_\star\mathbf{u}';\mathbf{u}''$. Indeed, since $\equiv_\star\subseteq\equiv$ and $u\neq u'$, the equality could hold if and only if $u^+=u'/u$ and $u''=u/u'$. But we reject also this case because, due to the distributivity of u and u', both $\rho_{(1)}$ and $\sigma_{(1)}$ could be lengthened, contrarily to the hypothesis. Hence, let $\zeta_1=\mathbf{u};\mathbf{u}^+$, $\zeta_2=\mathbf{u}';\mathbf{u}''$ and consider the sublattice of $\rho\Downarrow^\star$ whose elements are $\rho_{(1)}$, $\rho_{(1)};\mathbf{u}$, $\rho_{(1)};\mathbf{u}'$, $\rho_{(1)};\zeta_1$,

$\rho_{(1)}; \zeta_2$ and $\rho_{(1)}; (\zeta_1 \sqcup \zeta_2)$. This sublattice is clearly not distributive, therefore $\rho \Downarrow^\star$ is not distributive. ■

Notice that it is possible to find counterexamples to the Theorem 4.2 if one between the conditions about $\equiv_\star$ is removed. Indeed, if $\equiv_\star$ does not contain $\equiv^\diamond$ then take the least congruence which equates the derivations $(\lambda x.\, y)(\mathbf{I}z) \xrightarrow{\varepsilon} y$ and $(\lambda x.\, y)(\mathbf{I}z) \xrightarrow{2} (\lambda x.\, y)z \xrightarrow{\varepsilon}$. It gives prime algebraic partial orders as derivation spaces but it is not contained into $\equiv_d$. Similarly if we maintain the constraint $\equiv^\diamond \subseteq \equiv_\star$ and remove the congruence requirement, the least equivalence equating the above two derivations and containing $\equiv^\diamond$ is not contained into $\equiv_d$.

5 Turning a normalizing reduction strategy into a distributive one

We have started our investigations about concurrent semantics of the λ-calculus in order to formalize the encoding of evaluation mechanisms of a λ-term into a concurrent process (a prime event structure). This has led to a new semantics: distributive permutation equivalence. We also argue that call-by-value can be trivially encoded by a concurrent process because the derivation spaces it shows up are distributive lattices. Anyway, the call-by-value reduction strategy has the following drawbacks:

1. its normal forms have still β-redexes (example $\lambda x.\, \Omega$);

2. it is not normalizing (example $(\lambda x.\, y)\Omega$).

We can overcome the above drawbacks by investigating the transformation of a normalizing one-step reduction strategies into a distributive one which reduces as many redexes as the sequential-one. Let $\mathcal{N}$ be a reduction strategy, in the following $\mathcal{O}_\mathcal{N}(t)$ will denote the set of redexes offered by $\mathcal{N}$ at t. A one-step reduction strategy is such that $\#(\mathcal{O}_\mathcal{N}(t)) \leq 1$; furthermore, in this case, we denote with $\mathbf{red}^n_\mathcal{N}(t)$ the derivation $\mathbf{u}_1; \cdots \mathbf{u}_n$ where $t_1 = t$ and $u_i = \mathcal{O}_\mathcal{N}(t_i)$. The problem we tackle here can be formalized as follows:

> given a one-step normalizing reduction strategy $\mathcal{N}$, turn it into a *maximally* distributive one $\mathcal{N}_d$, i.e. a strategy offering, for every term t, a distributive set of redexes such that $\mathcal{O}_\mathcal{N}(t) \in \mathcal{O}_{\mathcal{N}_d}(t)$ and, for every n, $\mathbf{red}^n_\mathcal{N}(t)$ is a derivation of $\mathcal{N}_d$, too.

Notice that if the above transformation were possible we could safely encode the reduction strategy $\mathcal{N}_d$ into an event structure and, therefore, execute it concurrently.

Below we need of the notion of *context*, already defined in Section 2.

Definition 5.1 Let C be a context and let u_h be the occurrence number of its hole. Assume that C has a redex u such that $u|u_h$ or $u \prec u_h$. A one-step normalizing reduction strategy is *structurally unfair* w.r.t. C if for every λ-term t, the reduction strategy, when evaluating $C[t]$, reduces the redex at u if and only if t is in normal form. ■

Theorem 5.2 Turning a structurally unfair normalizing one-step reduction strategy into a maximally distributive one is not recursive.

Proof: Let $\mathcal{N}$ be a structurally unfair normalizing one-step reduction strategy, C be the context whose existence is ensured by the hypothesis and u be the occurrence number of the redex inside C. By contradiction, assume the existence of a recursive transformation $\Im$ which gives the maximally distributive strategy w.r.t. $\mathcal{N}$. Then, given $u' \in \mathcal{O}_\beta(t)$ it is recursive if $u' \in \mathcal{O}_{\Im(\mathcal{N})}(t)$ or not. Then, consider the following function:

$$\Im'(t) = \begin{cases} 1 & \text{if} \quad u \in \mathcal{O}_{\Im(\mathcal{N})}(C[t]) \\ 0 & \text{if} \quad u \notin \mathcal{O}_{\Im(\mathcal{N})}(C[t]) \end{cases}$$

Observe that the hypothesis on $\Im$ guarantees that $\Im'$ is a recursive decision function for the set $\{t \mid t \text{ has a normal form }\}$. But this is false because such a function should be recursive enumerable but not recursive. ■

Corollary 5.3 The leftmost-outermost reduction strategy can not be turned effectively into a maximally distributive one.

Proof: It is enough to show that leftmost-outermost is structurally unfair. To this aim consider the context $(x\,[\,])(\mathbf{I}x)$. ■

6 Conclusion

In this paper we have defined distributive equivalence, i.e. the coarsest congruence over λ-calculus derivations among those

1. that are contained into permutation equivalence and permute at least disjoint reductions;

2. yielding derivation spaces which are prime algebraic cpo's.

We have also shown that finding the maximal distributivity into a sequential reduction strategy is generally not recursive. Here we try to generalize our results to term rewriting systems. Since distributive permutation equivalence has been defined in an independent way from the language, its definition holds for the more general case of term rewriting systems, too. The theorem that the downward closure of derivations w.r.t. the distributive embedding is a distributive lattice uses the function of residual between derivations which is total only for orthogonal term rewriting systems. Anyway we claim that, in general, the function of residual is total between derivations which are compatible in the distributive derivation spaces. Therefore prime algebraicity of the derivation spaces should turn out with proofs which are similar to those in Section 3. On the other hand, the coarsest equivalence theorem should be more difficult to generalize. Indeed, in Theorem 4.2 we use a property of λ-calculus, which, given a term t with a redex, allows us to built up a new term having t as a subterm and having another redex disjoint from those in t. Of course, for every term rewriting system having the same property, Theorem 4.2 still holds.

A problem which we leave open concern the formal comparison between the encoding of an evaluation strategy raised from this paper and those which can be given by using an higher order process algebra, like the π-calculus [9]. We feel that, as soon as π-calculus is equipped with an event structure semantics, if $\mathcal{E}$ is the event structure yielded by the π-calculus encoding, then $\mathcal{E}$ should be obtained through an operation of *refinement* from the event structure raising from our encoding. Refinement is due to some book-keeping needed by the process algebra in order to encode β-reductions.

Acknowledgements: The author is very indebted with Miranda Mowbray for her meticulous job on an earlier draft of this paper which improved correctness and readability. Special thanks to Ugo Montanari for the stimulating discussions that led to this paper.

References

[1] G. Boudol. Computational semantics of term rewriting systems. In M. Nivat and J. C. Reinolds, editors, *Algebraic Methods in Semantics*, pages 169 – 236. Cambridge University Press, 1985.

[2] S. Burris and H.P. Sankappanavar. *A Course in Universal Algebra*. Springer-Verlag, 1981.

[3] H.B. Curry and R. Feys. *Combinatory Logic*, volume 1. North-Holland, 1974.

[4] G. Huet and J.J. Lévy. Computations in orthogonal systems, 1. In J.L. Lassez and G. Plotkin, editors, *Computational Logic, Essays in honour of A. Robinson*, chapter 11. MIT Press, 1991.

[5] C. Laneve and U. Montanari. Axiomatizing permutation equivalence in the λ-calculus. In 3^{th} *Int. Conf. on Algebraic and Logic Programming*, volume 632 of *Lecture Notes in Computer Science*, pages 350 – 363. Springer-Verlag, 1992.

[6] J.J. Lévy. *Réductions correctes et optimales dans le lambda calcul*. PhD thesis, Université Paris VII, 1978.

[7] J.J. Lévy. Optimal reductions in the lambda-calculus. In J.P. Seldin and J.R. Hindley, editors, *To H.B. Curry, Essays on Combinatory Logic, Lambda Calculus and Formalism*, pages 159 – 191. Academic Press, 1980.

[8] J. Meseguer. Conditional rewriting systems. *Theoretical Computer Science*, 96(1):73 – 155, 1992.

[9] R. Milner. Functions as processes. In M. Paterson, editor, *Proceedings* 17^{th} *ICALP,* Warwick, volume 443 of *Lecture Notes in Computer Science*, pages 167–180. Springer-Verlag, July 1990.

[10] G. Winskel. An introduction to event structures. In *REX School/Workshop on Linear Time, Branching Time and Partial Order in Logics and Models for Concurrency,* Noordwijkerhout, volume 354 of *Lecture Notes in Computer Science*, pages 364–397. Springer-Verlag, 1989.