



HAL
open science

Confluence in (Un)Typed Higher-Order Theories I

Gaspard Ferey, Jean-Pierre Jouannaud

► **To cite this version:**

Gaspard Ferey, Jean-Pierre Jouannaud. Confluence in (Un)Typed Higher-Order Theories I. 2019. hal-02096540v2

HAL Id: hal-02096540

<https://hal.science/hal-02096540v2>

Preprint submitted on 5 Dec 2019 (v2), last revised 21 Jan 2020 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Confluence in UnTyped Higher-Order Theories: Part I

GASPARD FÉREY AND JEAN-PIERRE JOUANNAUD

User-defined higher-order rewrite rules are becoming a standard in proof assistants based on intuitionistic type theory. This raises the question of proving that they preserve the properties of beta-reductions for the corresponding type systems. In a series of papers, we develop techniques based on van Oostrom's decreasing diagrams that reduce confluence proofs to the checking of various forms of critical pairs for higher-order rewrite rules extending beta-reduction on pure lambda-terms. The present paper concentrates on the case where rewrite rules are left-linear and critical pairs can be joined without using beta-rewrite steps. The case where beta-rewrite steps are needed will be considered in the next paper. The case of non-left-linear rules will be addressed in the third paper. The application of these results to various encodings of type theories into the logical framework LF extended with rewrite rules will be carried out in the last paper.

Additional Key Words and Phrases: Lambda calculus, Church-Rosser property, Confluence, Decreasing diagrams, Critical pairs

ACM Reference format:

Gaspard Férey and Jean-Pierre Jouannaud. 2017. Confluence in UnTyped Higher-Order Theories: Part I. *Proc. ACM Program. Lang.* 1, 1, Article 1 (January 2017), 29 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 INTRODUCTION

The two essential properties of a type theory, consistency and decidability of type checking, follow from three simpler ones: type preservation, strong normalization and confluence. In dependent type theories however, confluence is often needed to prove type preservation and strong normalization, making all three properties interdependent if termination is used in the confluence proof. This circularity can be broken in two ways: by proving all properties together within a single induction [7]; or by proving confluence on untyped terms first, and then successively type preservation, confluence on typed terms, and strong normalization. We develop the latter way here, focusing on untyped confluence.

In Coq and Agda, rewrite rules introduced by the user originate from the definition of inductive types of some form. They satisfy a precise format which has been well studied, for which confluence is always satisfied. But Agda and Coq developers have recently announced the development of new versions that would allow user-defined rewrite rules [3], as is already the case in DEDUKTI [5], and several on-going developments in Agda are already using this facility. Investigating the preservation of confluence by user-defined rewrite rules in λ -calculus appears therefore to be very timely.

DEDUKTI has been mostly used so far as a logical framework, user-defined rewrite rules originating then from complex higher-order encodings for which inductive types do not provide enough flexibility. Let \mathcal{R} be the set of user-defined rewrite rules, which come in addition to the β -rule. The rewrite relation underlying the type theory is then generated by both \mathcal{R} and the β -rule. Studying the meta-theory of such a type theory implies investigating the confluence property of $\beta \cup \mathcal{R}$.

There are three main tools for analyzing confluence of a rewrite relation: Newman's Lemma [14], Hindley-Rosen's Lemma [9], and van Oostrom's Theorem which generalizes both previous ones [19]. Since beta rewrites are non-terminating in pure lambda calculus, Newman's Lemma does not

2017. 2475-1421/2017/1-ART1 \$15.00
<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

50 apply. And if the rules have non-trivial critical pairs, then Hindley-Rosen's Lemma does not apply
 51 either. Even the subtle use of Hindley-Rosen's Lemma allowing for development-closed critical
 52 pairs [18] is too restrictive for practical usage. The way out is the use of van Oostrom's decreasing
 53 diagrams [17]. The fact that beta reductions do not terminate for pure lambda terms is no obstacle
 54 since we do not rely on termination for showing confluence when using decreasing diagrams. A
 55 further reason for considering pure lambda terms is that it is then easy to deduce confluence for
 56 any type system, including dependent type systems, for which the rules enjoy type preservation.

57 Van Oostrom's theorem is abstract, its application to term rewriting relations conceals many
 58 difficulties. Further, neither confluence nor termination are preserved by adding a confluent and
 59 terminating set of rewrite rules to a λ -calculus. A counter-example to termination in the simply
 60 typed λ -calculus is given in [15]. Numerous counter-examples to confluence in the pure λ -calculus
 61 are given in [11]. The problem we address is by no means simple.

62 Our untyped λ -calculus is intended to fit with the implementation of DEDUKTI. The format of
 63 rules is classical: left-hand sides must be patterns [12, 13], which are extremely useful for describing
 64 encodings of a type theory in another, a keen application to us. Considering untyped terms requires
 65 simple adaptations of the usual higher-order rewriting definitions. In particular, we shall consider
 66 that the meta-variables used in rules have an arity which is not fixed, but bound, hence allowing for
 67 implicit arguments. These adaptations impact unification: we shall precisely analyze unification of
 68 linear untyped patterns and show the existence of most general unifiers computable in linear time.

69 Our contribution is that a set $\mathcal{R}ll$ of rules which is terminating on the set of pure λ -terms and
 70 whose left-hand sides are linear patterns, preserves confluence of the λ -calculus if its critical pairs
 71 are joinable by using rules in $\mathcal{R}ll$.

72 This result is then demonstrated with the example of a theory of global states due to Plotkin and
 73 Power [16], whose rules have overlapping linear higher-order patterns as left-hand sides. As we
 74 shall see, its critical pairs are not development closed. The confluence of this example had indeed
 75 been shown already in [8]. Hamana shows first that the (simply) typed rules are terminating, then
 76 that the higher-order critical pairs are joinable, using Newman's Lemma to deduce its confluence.
 77 Our methods apply independently of the typing system, hence we can deduce that the same example
 78 remains confluent when using a dependently typed discipline.

79 We recall the notion of labeled reduction and decreasing diagram in Section 2, and describe our
 80 higher-order setting in Section 3. Matching and unification of untyped patterns and basic properties
 81 of untyped higher-order rewriting is considered in Section 4. Local rewriting peaks are analyzed in
 82 Section 5. Our confluence result is stated and proved in Section 6. Significance of the framework,
 83 and of our result, is discussed in Conclusion.

84 2 LABELED REDUCTIONS

85 2.1 Reductions

86 Given a binary relation \longrightarrow on terms, called *rewriting*, we use: \longleftarrow for its inverse, \Longrightarrow for its
 87 parallelization, allowing one to rewrite at once several subterms of a given term, when none is a
 88 subterm of another, and \longleftrightarrow , \longrightarrow , and \longleftarrow , for its closures by, respectively, symmetry; reflexivity
 89 and transitivity (called *derivation*); and reflexivity, symmetry and transitivity (called *convertibility*).

90 A term s is in *normal form* if there is no t such that $s \longrightarrow t$. We define a *normal form* for an arbitrary
 91 term s as a term t in normal form, denoted by $s \downarrow$, such that $s \longrightarrow t$. Termination is the impossibility of
 92 an infinite rewriting sequence $t_0 \longrightarrow t_1 \longrightarrow \dots \longrightarrow t_n \longrightarrow \dots$. Termination guarantees the existence
 93 of normal forms for every term. A *local peak* is a triple of terms (s, u, t) such that $s \longleftarrow u \longrightarrow t$; u is
 94 the *source* and s, t are its *reducts*. Two terms s, t are *joinable* if $s \longrightarrow v \longleftarrow t$ for some v , making
 95
 96
 97
 98

the peak $s \leftarrow u \rightarrow t$ *joinable*. The property that every two convertible terms are joinable is called *confluence* (or *Church-Rosser*). Confluence guarantees the unicity of normal forms for every term.

When rewriting terminates, it is well-known that the joinability of all local peaks implies the confluence property, this is the so-called Newman's lemma. When it does not, it is then necessary to strengthen joinability, this is the rôle of *decreasing diagrams*.

2.2 Decreasing diagrams

In the following, we consider rewrite relations all of whose elementary steps are equipped with a label belonging to some well-founded set whose strict partial order is denoted by \triangleright .

Definition 2.1 (Decreasing diagram [17]). Given a labeled relation \rightarrow on an abstract set, we denote by $DS(m, n)$ the set of *decreasing rewrite sequences* of the form $u \xrightarrow{\delta} v$ or $u \xrightarrow{\gamma} s \xrightarrow{n} t \xrightarrow{\delta} v$ such that the labels in γ and δ are strictly smaller than, respectively, m , and, m or n . The steps labeled by γ , n and δ , are called the *side steps*, *facing step* and *middle steps* of the decreasing sequence, respectively.

Given a local peak $v \xleftarrow{m} u \xrightarrow{n} w$, a *decreasing (rewrite) diagram* is a pair of derivations from v and w to some common term t , belonging to $DS(m, n)$ and $DS(n, m)$, respectively.

Decreasing rewrite diagrams are represented at Figure 2 and abbreviated as DDs. Note that a facing step of a decreasing diagram may be missing, its side steps are then absorbed by the middle ones. A more general notion of decreasing diagram appears in [19], we won't need it here.

THEOREM 2.2 ([17]). *A labeled relation is Church-Rosser if all its local peaks have DDs.*

van Oostrom's theorem generalizes to rewriting modulo an equational theory, for example $=_{\alpha}$ in the λ -calculus, in which case \triangleright must be compatible with the equational theory [10]. This is of course true of $=_{\alpha}$ which must be built-in any definition of reduction over lambda terms. Further, equational steps with $=_{\alpha}$ must have a minimal label, which is easy to achieve.

3 HIGHER-ORDER REWRITING

Given an *untyped* lambda calculus generated by a vocabulary made of three pairwise disjoint sets, a signature \mathcal{F} of *function symbols*, a set \mathcal{X} of *variables*, and a set \mathcal{Z} of *meta-variables*, we are interested in the calculus $\lambda\mathcal{F}$, whose reduction relation extends the β -rule of the underlying λ -calculus by a set R of user-defined rewrite rules built over that vocabulary. Were we to analyze the confluence of R alone, then, the situation would be similar to the first-order case, at least when left-hand sides of rules are patterns [12]. Unfortunately, confluence of $R \cup \beta$ cannot, in general, be deduced from the confluence of its two components.

3.1 Terms in $\lambda\mathcal{F}$

$\lambda\mathcal{F}$ is a mix of the pure lambda-calculus and Klop's combinatory reduction systems [11], that fits with DEDUKTI [5]. Terms are those of the pure lambda calculus enriched with \mathcal{F} -headed terms of the form $f(\bar{u})$ with $f \in \mathcal{F}$ and *meta-terms* of the form $Z[\bar{v}]$ with $Z \in \mathcal{Z}$. Only variables can be abstracted over. Elements of the vocabulary have arities, denoted by vertical bars as in $|a|$. Variables have arity zero. The grammar of terms is the following:

$$u, v := x \in \mathcal{X} \mid (u \ v) \mid \lambda x. u \mid f(\bar{u}) \mid Z[\bar{v}] \quad \text{where } f \in \mathcal{F}, |\bar{u}| = |f|, Z \in \mathcal{Z} \text{ and } |\bar{v}| \leq |Z|$$

Following usage, we don't duplicate parentheses, writing $f(x \ y)$ for $f((x \ y))$, and use brackets instead of parentheses for meta-variables. It is sometimes convenient to name the head symbol

of the expression ($s \ t$): we use the symbol $@$ for that purpose throughout the paper. We use the small letters f, g, h, \dots for function symbols and x, y, z, \dots for variables, and reserve capital letters X, Y, Z, \dots for meta-variables. When convenient, a small letter like x may denote any variable in $\mathcal{X} \cup \mathcal{Z}$. We use the notation $|_$ to denote various quantities besides arities, such as the length of a list, the size of an expression or the cardinality of a set. Given a list \bar{u} , $\bar{u}[m..n]$ denotes the finite sublist u_m, \dots, u_n . The list \bar{u} is omitted in case it is the list of natural numbers.

Unlike function symbols and Klop's meta-variables, meta-variables here have an arity which is not fixed, but bounded, a handy feature used in DEDUKTI that allows meta-variables to take *implicit arguments*. This peculiarity has several objectives. First-of-all, the number of dependent arrows in a dependent type T is not fixed, it may depend upon, say, the value of a natural number this type depends upon. However, any occurrence of a meta-variable of type T used in a rewrite rule must have a finite number of arguments, the maximum of these numbers can then be taken as the arity of that meta-variable. Another use of this facility in DEDUKTI is to speed up computations by avoiding type-checking terms along rewriting derivations. The pattern matching algorithm, as we shall see in Section 4, requires using the arity of meta-variables instead of their type. Finally, verifying in DEDUKTI that rewrite rules preserve types is based, as expected, on solving type equality constraints, which in turn requires inferring the arities of the meta-variables that occur in those rules.

Positions in terms are words over the natural numbers (assuming $|\lambda x._| = 1$), using \cdot for concatenation, Λ for the empty word, P/p for $\{q : p \cdot q \in P\}$, \leq_p for the prefix order (*above*), \geq_p for its inverse (*below*), $>_p$ for the strict part of \geq_p , and $p\#q$ for $\neg(>_p \vee \leq_p)$ (*parallel*). An order $>$ on positions is extended to sets of positions as follows: $P > Q$ iff $\forall p \in P \exists q \in Q$ such that $p > q$.

Given a term M , we use $\mathcal{P}os(M)$, $\mathcal{V}\mathcal{P}os(M)$, $\mathcal{M}\mathcal{P}os(M)$ for the following respective sets of positions of M : all positions, the positions of free variables, and of meta-variables, and $\mathcal{V}ar(M)$ and $\mathcal{M}\mathcal{V}ar(M)$ for its sets of free variables and of meta-variables respectively. A term M is *ground* if $\mathcal{V}ar(M) = \emptyset$, *closed* if $\mathcal{M}\mathcal{V}ar(M) = \emptyset$, and *linear* if $|\mathcal{M}\mathcal{P}os(M)| = |\mathcal{M}\mathcal{V}ar(M)|$. Given now a term M and a position $p \in \mathcal{P}os(M)$, we use $M(p)$ for its symbol at positions p , $M|_p$ for the subterm of M at position p , $M[\bar{N}]_p$ for the term obtained by replacing in M the subterm $M|_p$ by the term N . The latter notations extends to sets P of parallel positions in $M[\bar{N}]_P$ or $M[\bar{N}]_P$ in case all terms in \bar{N} are identical to the term N . This use of brackets in the meta-language of terms is reminiscent of its use in the term language, namely in $Z[\bar{v}]$. Both kinds of brackets may occur in a same expression, as long as the replacement bracket is indexed by a position or set of positions. We sometimes use the notation $u[\bar{v}]_p$ to stipulate that the subterm of u at position p is the term v . The context is supposed to help discriminating between these different uses of the bracket notation.

A *substitution* is a *capture-avoiding* homomorphism written as $\sigma = \{x_1 \mapsto M_1, \dots, x_n \mapsto M_n\}$, or $\sigma = \{\bar{x} \mapsto \bar{M}\}$, where $M_i = \lambda \bar{y}_i. N_i$ with $|\bar{y}_i| \geq |x_i|$. Note that x_i denotes here an element of $\mathcal{X} \cup \mathcal{Z}$, hence its arity may be non-null. $\mathcal{D}om(\sigma) = \{x_1, \dots, x_n\} \subseteq \mathcal{X} \cup \mathcal{Z}$ is the *domain* of σ while $\mathcal{R}an(\sigma) = \bigcup_{i=1}^n \mathcal{V}ar(M_i)$ is its *image*. The substitution σ is *ground* (resp., *closed*) when so are all M_i 's. A substitution σ can be *restricted to* or *deprived from* (meta-)variables in some set V , written $\sigma|_V$ and $\sigma_{\setminus V}$ respectively. As in λ -calculus, substituting in terms requires renaming bound variables to avoid capturing free ones. Using post-fixed notation, $x_i\sigma = t_i$ and $y\sigma = y$ if $y \notin \mathcal{D}om(\sigma)$; $f(\bar{t})\sigma = f(\bar{t}\sigma)$; $(u \ v)\sigma = (u\sigma \ v\sigma)$; and $(\lambda x.u)\sigma = \lambda x.u\sigma$ if $x \notin \mathcal{D}om(\sigma) \cup \mathcal{R}an(\sigma)$ (otherwise, as announced, x must be renamed away from $\mathcal{D}om(\sigma) \cup \mathcal{R}an(\sigma)$ in $\lambda x.u$.) The additional rule for meta-variables is as follows: if $Z \mapsto \lambda \bar{x}.s \in \sigma$, then $Z[\bar{u}]\sigma = \lambda \bar{x}[m+1..n].s\{\bar{x}[1..m] \mapsto \bar{u}\sigma\}$, where $|\bar{u}| = m \leq n = |Z|$, hence delaying the replacement of those arguments of Z that are missing. The result $t\sigma$ of substituting the term t is called an *instance* (of t) and the operation itself an *instantiation*.

Arities extend naturally to terms, writing $ar(t)$ for the *arity* of a term t , by induction on their structure: $ar(\lambda x.t) = 1 + ar(t)$, $ar(X[\bar{t}]) = |X| - |\bar{t}|$ and $ar((u \ v)) = ar(x) = ar(f(\bar{u})) = 0$.

Substitution of meta-variables was introduced by Klop in the case of a fixed arity [11]. Our definition ensures the straightforward properties that arities are non-decreasing under substitution, hence provide enough abstractions for all meta-terms of the form $X[\bar{u}]$ encountered in a computation. (So does of course Klop's definition for the case of fixed arities.)

Example 3.1. Let X be a ternary meta-variable, $s = \lambda x.(X[x, f(x, y)] g(y))$ and $\sigma = \{X \mapsto \lambda xzz'.h(z', z), y \mapsto h(a, a)\}$. Then, $s\sigma = \lambda x.(\lambda z'.h(z', f(x, h(a, a)))) g(h(a, a))$. Then, $ar(s) = 1 + (3 - 2) = 2 = ar(s\sigma)$.

Rewriting terms extends to substitutions as expected, while substitutions are extended to sequences of terms and to substitutions in the natural way, keeping the same post-fixed notation.

LEMMA 3.2. *Given $u, \sigma, \tau, (u\sigma)\tau = u(\sigma\tau)$ (we write $u\sigma\tau$).*

Given a term u and a list $P = \{p_i\}_{i=1}^n$ of parallel positions in u , we define the term obtained by *splitting* u along P as $\underline{u}_P = u[Z_1[\bar{x}_1]]_{p_1} \dots [Z_n[\bar{x}_n]]_{p_n}$ (u is cut below P) and its associated substitution by $\bar{u}^P = \{Z_i \mapsto \lambda \bar{x}_i.u|_{p_i}\}_{i=1}^n$ (u is cut above P), where, for all $i \in [1, n]$, \bar{x}_i is the list of all variables of $u|_{p_i}$ bound in u above p_i and Z_i is a fresh meta-variable of arity (exactly) $|\bar{x}_i|$. The definition of substitution for meta-variables ensures that $\underline{u}_P \bar{u}^P = u$, which justifies writing $u = u|_P$ as a familiar shorthand. Note the two kinds of brackets in \underline{u}_P .

Our use of splitting in this paper will be systematic unless it alters readability for no good reason. This invention permitted by Klop's notion of meta-variable, is the only technique we know of which allows to manipulate terms with binders safely, in case renaming of variables needs to take place independently in a term and in its context, as will often be the case here.

3.2 Functional reductions

Arrow signs used for rewriting will often be decorated, below by a name, and above by a position p or set of positions P , as in $s \xrightarrow[p]{R} t$ or by a property that this position or set of positions satisfies, as in $u \xrightarrow[R]{\geq p P} v$ and in $u = v \downarrow_{R}^{\geq p P}$ (u is obtained from v by normalizing its subterms $v|_{p \in P}$ with R).

Two different kinds of reductions coexist in $\lambda\mathcal{F}$, functional and higher-order reductions. Both are meant to operate on closed terms. However, rewriting open terms will sometimes be needed, in which case rewriting is intended to rewrite all their closed instances at once.

Functional reduction is the relation on terms generated by the rule $(\lambda x.u v) \xrightarrow[\beta_\alpha]{\beta} u\{x \mapsto v\}$. The usually omitted α -index stresses that renaming bound variables, called α -conversion, is built-in.

As is customary [13], the particular case for which v is a variable is denoted by β^0 . Note that instantiating a β^0 -step may yield a full β -step. For example, $s = (\lambda x.(\lambda y.g(y) x) a) \xrightarrow[\beta^0]{1,1} (\lambda x.g(x) a) \xrightarrow[\beta]{\Lambda} g(a)$

while $s \xrightarrow[\beta]{\Lambda} (\lambda y.g(y) a) \xrightarrow[\beta]{\Lambda} g(a)$. This is our main motivation for using Klop's notion of substitution for meta-variables, whose benefits will appear in the next subsection.

We will also use a particular case of extensionality, for meta-variables only: $\lambda z.X[\bar{u}, z] =_{M\eta} X[\bar{u}]$ if $|X| > |\bar{u}|$, z fresh. When oriented from left to right, $M\eta$ is terminating and confluent. It has an even more important property: assume σ is a substitution replacing X by $\lambda \bar{x}z.v$. Then $\lambda z.X[\bar{u}, z]\sigma = \lambda z.v\{\bar{x} \mapsto \bar{u}\} = X[\bar{u}]\sigma$. So, $M\eta$ -steps disappear when taking instantiations, a key property for us.

3.3 Higher-order reductions

Higher-order reductions result from rules whose left-hand sides are higher-order patterns in Miller's or Nipkow's sense [12], although they need not be typed:

Definition 3.3 (Pattern). A *pre-redex* of arity n in a term L is an unapplied meta-term $Z[\bar{x}]$ whose arguments \bar{x} are n pairwise distinct variables. A *pre-pattern* is a ground β -normal term all of whose meta-variables occur in pre-redexes. A *pattern* is a pre-pattern which is neither an abstraction nor a pre-redex.

It is important to assume, as does Nipkow, that pre-patterns are β -normal. Note that erasing types from a Nipkow's pattern yields a pattern in our sense, since his pre-redexes being of base type, they cannot be applied. This restriction isn't important until later, when we address the question of matching and unification of patterns.

The properties investigated below are true of pre-patterns, not only of patterns.

Observe that pre-redexes in pre-patterns can only occur at parallel positions, whose set plays a key rôle:

Definition 3.4 (Fringe). The *fringe* F_L of a pre-pattern L is the set of parallel positions of its pre-redexes. We denote by $\mathcal{FPos}(L) = \{p \in \mathcal{Pos}(L) : p <_{\mathcal{P}} F_L\}$ the (non-empty) set of *functional positions* of the pre-pattern L , and by $\mathcal{MVar}(L, o)$, for $o \in F_L$, the meta-variable Z such that $L|_o = Z[\bar{x}]$. We also define $F_\beta = \{1, 2\}$ for convenience.

Example 3.5. The term $L = f(\lambda xyz.g(X[x, y, z], X[x, y]))$ is a pattern. Its pre-redexes are the terms $X[x, y, z]$ and $X[x, y]$. Its fringe is the set $F_L = \{1^5, 1^4 \cdot 2\}$. The term $(f(\lambda xyz.g(X[x, y, z]))(a X))$ is also a pattern, its fringe is the set $\{1^6, 2^2\}$. Terms $f(\lambda x.X[x, x])$, $f(X[a])$, $f(X[Y])$, and $f(X Y)$, are no patterns.

Note that the set of functional positions coincides with the usual notion for first-order terms. Since patterns are ground terms, we have:

LEMMA 3.6. *Given a pre-pattern L , let $p \in F_L$ and $L|_p = Z[\bar{x}]$ be a pre-redex. Then, all variables in \bar{x} are bound above p in L .*

We can now define higher-order rules and rewriting:

Definition 3.7 (Rule). A (higher-order) *rule* is a triple $i : L \rightarrow R$, whose (possibly omitted) *index* i is a natural number, left-hand side L is a pattern, and right-hand side R is a ground β -normal term such that $\mathcal{MVar}(R) \subseteq \mathcal{MVar}(L)$. The rule is *left-linear* if L is linear and *right-linear* if R is linear.

So, rules are pairs of (specific) ground terms. They may have meta-variables, but don't admit free variables. This allows to clearly separate the object language (which has no meta-variables), from the meta-language (which has meta-variables). Rules, critical pairs and splittings belong to the meta-language, which serves analyzing the properties of the language. The role taken by free variables in first-order rules is therefore taken here by meta-variables of arity zero.

Definition 3.8 (Higher-order untyped rewriting). Given an open term u , a position $p \in \mathcal{Pos}(u)$, and a rule $i : L \rightarrow R$ then u rewrites with i at p , written $u \xrightarrow[p]{i} v$, iff $u|_p = L\gamma$ for some substitution γ , and $v = u[X[\bar{x}]]_p \{X \mapsto \lambda \bar{x}. R\gamma\} = u[R\gamma]_p$, where \bar{x} is the list of variables of $u|_p$ which are bound above the position p in u . We write $u \xrightarrow[\mathcal{R}]{p} v$ for $\exists i \in \mathcal{R}. u \xrightarrow[p]{i} v$.

Let's now make our splitting notations fully explicit. Whenever $u \xrightarrow[p]{i} v$, we have by definition:

- $\underline{u}_p = u[X[\bar{x}]]_p$ and $\bar{u}^p = \{X \mapsto \lambda \bar{x}. u|_p\}$ with \bar{x} variables bound above p in u
- $u = \underline{u}_p \bar{u}^p = \underline{u}_p \{X \mapsto \lambda \bar{x}. u|_p\} = \underline{u}_p \{X \mapsto \lambda \bar{x}. L\gamma\}$
- $v = \underline{u}_p \{X \mapsto \lambda \bar{x}. R\gamma\}$, hence $\underline{v}_p = \underline{u}_p$, $\bar{v}^p = \{X \mapsto \lambda \bar{x}. R\gamma\}$ and $v|_p = R\gamma$.

Example 3.9. Let $L = \text{der}(\lambda x. \text{sin}(F[x])) \rightarrow R = \lambda x. \text{cos}(F[x])$, and take for σ the identity substitution $\{F \mapsto \lambda x. x\}$. Then, $L\sigma = \text{der}(\text{sin}(x))$ and $R\sigma = \text{cos}(x)$, hence $\text{der}(\text{sin}(x)) \rightarrow \text{cos}(x)$.

In sharp contrast with Nipkow [12], we observe that we do not need matching modulo β^0 , since the corresponding β^0 -steps are now hidden in the Klop's definition of substitution for meta-variables. We will however show that our main confluence result applies to Nipkow's definition: the use of Klop's definition of substitution for meta variable can be seen as a technical choice.

Besides, we do *not* assume that u , or v , is β -normal, or even β -normal up to position p . We cannot for two reasons: β -normal forms may not exist, and we need monotonicity and stability properties:

LEMMA 3.10 (SPLITTING ABOVE). Let $s \xrightarrow{q}_{L \rightarrow R} t$. Then, $\bar{s}^q \xrightarrow{\sigma}_{L \rightarrow R}$ and $t = \underline{s}_q \sigma$.

LEMMA 3.11 (MONOTONICITY). Let $s \xrightarrow{p}_{L \rightarrow R} t$ and u a term such that $q \in \text{Pos}(u)$. Then, $u[s]_q \xrightarrow{q \cdot p}_{L \rightarrow R} u[t]_q$.

Monotonicity follows directly from definition and Lemma 3.10. Stability is just as easy.

LEMMA 3.12 (STABILITY). Let $s \xrightarrow{p}_{L \rightarrow R} t$ and σ a substitution. Then $s\sigma \xrightarrow{p}_{L \rightarrow R} t\sigma$.

PROOF. By definition of higher-order rewriting, $s|_p = L\gamma$ for some substitution γ , and $t = s[R\gamma]_p$. We have $s\sigma|_p = s|_p\sigma = L\gamma\sigma$ and $t\sigma = s[R\gamma]_p\sigma = s\sigma[R\gamma\sigma]_p$ yielding the result. \square

LEMMA 3.13 (SUBSTITUTION LEMMA). Let $u \xrightarrow{p}_{\mathcal{R}II} v$ and $\sigma \xrightarrow{\tau}_{\mathcal{R}II}$. Then, $u\sigma \xrightarrow{v\tau}_{\mathcal{R}II}$.

PROOF. We first prove $u\sigma \xrightarrow{\tau}_{\mathcal{R}II} u\tau$ by induction on u :

- $u = f(\bar{u})$ with $f \in \mathcal{V} \cup \{\@, \lambda\}$ By induction hypothesis, $\bar{u}\sigma \rightarrow \bar{u}\tau$. By monotonicity, $f(\bar{u}\sigma) \rightarrow f(\bar{u}\tau)$. Conclusion follows.
- $u = x$. This case is straightforward.
- $u = X[\bar{u}]$ with $X \notin \text{Dom}(\sigma)$. Similar to the first case.
- $u = X[\bar{u}]$ with $X\sigma = \lambda \bar{x}. w$, hence $w \xrightarrow{\mathcal{R}II} w'$ and $X\tau = \lambda \bar{x}. w'$. By induction hypothesis, $\bar{u}\sigma = \bar{u}\tau$. Hence $u\sigma = w\{\bar{x} \mapsto \bar{u}\sigma\} \rightarrow w\{\bar{x} \mapsto \bar{u}\tau\}$ (by monotonicity) $\rightarrow w'\{\bar{x} \mapsto \bar{u}\tau\}$ (by stability) $= u\tau$.

Since $u\tau \xrightarrow{v\tau}_{\mathcal{R}II}$ by stability, we conclude that $u\sigma \xrightarrow{v\tau}_{\mathcal{R}II}$. \square

3.4 Rewrite theories

Definition 3.14. A $\lambda\mathcal{F}$ -rewrite theory is a pair $(\mathcal{F}, \mathcal{R})$ made of a user's signature \mathcal{F} and a set \mathcal{R} of higher-order rewrite rules on that signature, defining the rewrite relation $\xrightarrow{\lambda\mathcal{F}}$ of $\lambda\mathcal{F}$ as $\xrightarrow{\mathcal{R} \cup \beta_\alpha}$.

Rewrite theories are used in DEDUKTI [1] to define the conversion rule of the calculus, which is, as is customary, untyped. The rewrite relation implemented in DEDUKTI is indeed the one we just described, Klop's notion of substitution for meta-variables being implemented via a priority mechanism.

The main question addressed in this paper is whether a $\lambda\mathcal{F}$ -rewrite theory is confluent (Church-Rosser), and how to show its confluence by calculating and inspecting critical pairs of some form. We shall focus on rewrite theories for which the set of rules \mathcal{R} satisfies linearity assumptions. We say that $\lambda\mathcal{F}$ is : a *left-linear* theory if \mathcal{R} is a set of left-linear rules; a *right-linear* theory if \mathcal{R} is a set of right-linear rules; a *semi-linear* theory if \mathcal{R} is made of rules which are of either kind.

In this paper, we restrict our attention to left-linear rewrite theories $(\mathcal{F}, \mathcal{R}II)$.

3.5 The rewrite theory of global states

Our running example here will be Plotkin's and Power's theory of global states for a single location [16]. It is described by two types, Val for values and A for states, a unary operation lk for looking up a state, a binary operation ud for updating a state, and five higher-order rules which satisfy our format, the meta-variables having arities (unlike in the original article). First, the signature:

$$lk : (Val \rightarrow A) \rightarrow A \quad | \quad ud : Val, A \rightarrow A$$

$lk(\lambda v.t)$ looks up the state, binds its value to v , and continues with t while $ud(v, t)$ updates the state to v , and continues with t . Types are given for a better understanding, they do not play any role here. Let us now give the rules, using X (resp. Y) (resp. U, V, W) for meta variables of arity 1 (resp. 2) (resp. 0). We also use Z , whose arity will have to be given. These meta-variables may appear primed when too many of a given arity are needed, as it will be the case when computing critical pairs.

$$\begin{array}{ll}
 (ll) & lk(\lambda w.lk(\lambda v.Y[w, v])) \rightarrow lk(\lambda v.Y[v, v]) \quad (ll) \\
 (lu) & lk(\lambda v.ud(v, X[v])) \rightarrow lk(\lambda v.X[v]) \quad | \quad lk(\lambda v.U) \rightarrow U \quad (l) \\
 (ul) & ud(V, lk(\lambda v.X[v])) \rightarrow ud(V, X[V]) \quad | \quad ud(U, ud(V, W)) \rightarrow ud(V, W) \quad (uu)
 \end{array}$$

This typed higher-order theory was studied by Hamana, who was interested in its confluence investigated with his Haskell-based analysis tool SOL [8]. Our presentation is a simplification of Hamana's, whose one rule was actually superfluous. Note that all rules are left-linear.

In this example, all meta-variables take a constant number of arguments, equal to their arity. Using our meta-variables with a bounded arity, we can reformulate this system by eliminating its η -expansions:

$$\begin{array}{ll}
 (ll) & lk(\lambda w.lk(Y[w])) \rightarrow lk(\lambda v.Y[v, v]) \quad (ll) \\
 (lu) & lk(\lambda v.ud(v, X[v])) \rightarrow lk(X) \quad | \quad lk(\lambda v.U) \rightarrow U \quad (l) \\
 (ul) & ud(V, lk(X)) \rightarrow ud(V, X[V]) \quad | \quad ud(U, ud(V, W)) \rightarrow ud(V, W) \quad (uu)
 \end{array}$$

We could of course, eliminate the η -expansions from the left-hand sides, and keep them in the right-hand sides. We will see that the precise formulation of the rules, when there are many possible variations, impacts their confluence properties.

4 PATTERN MATCHING AND UNIFICATION OF LINEAR PATTERNS

Firing rules requires pattern matching an arbitrary term against a pattern, while computing critical pairs, which play a key role in the analysis of overlapping peaks, requires unifying two patterns. Both algorithms are described by rewrite rules operating on equational problems.

Definition 4.1. A (matching or unification) *equational problems* \mathcal{P} is a conjunction of elementary equations. An *elementary equation* is either the constant \perp or is of the form $u = v$ in which u is a pre-pattern, v is either a pre-pattern (unification case), or an arbitrary term (matching case).

We now define solutions and unifiers of an equational problem, the unifiers being representations of their solutions. It is important to note here that patterns have no free variables. This implies that solutions and unifiers of unification problems can be restricted to be ground, since additional variables are not needed for expressing unifiers, we can use meta-variables of arity zero instead.

Definition 4.2. A *solution* of a matching problem \mathcal{P} different from \perp is a substitution γ such that $Dom(\gamma) \subseteq \mathcal{Z}$ and for all equations $u = v \in \mathcal{P}$, $u\gamma =_\alpha v$.

A *solution* of a unification problem \mathcal{P} different from \perp is a closed, ground substitution γ such that $Dom(\gamma) \subseteq \mathcal{Z}$, and for all equations $u = v \in \mathcal{P}$, then $u\gamma =_\alpha v\gamma$.

393 A *unifier* of a unification problem \mathcal{P} different from \perp is a ground substitution γ such that
 394 $\text{Dom}(\gamma) \subseteq \mathcal{Z}$, and for all equations $u = v \in \mathcal{P}$, then $u'\gamma =_{\alpha, M\eta} v'\gamma$.

395 The constant \perp has no (matching or unification) solution nor unifier.

396 Unifiers equate terms of an equation modulo renaming, but also modulo extensionality for meta-
 397 variables. As we have seen, the latter steps will disappear by instantiation of the meta-variables: unifiers
 398 are not solutions but representations of solutions via their closed instances.

399 *Definition 4.3.* A unification problem \mathcal{P} is *linear* if no meta-variable occurs more than once in \mathcal{P} .
 400 A matching problem \mathcal{P} is *linear* if no meta-variable occurs more than once in the left-hand sides of
 401 the elementary equations of \mathcal{P} .
 402

403 In the sequel, we will usually omit $=_{\alpha}$, and also restrict ourselves to linear equational problems.
 404 Before to give the rules, we need the following preliminary definition:

405 *Definition 4.4.* A variable $x \in \mathcal{X}$ is *protected* in a pre-pattern u if all its occurrences in u belong
 406 to a pre-redex of u , that is, take place below F_u .
 407

408 For an example, x is protected in $f(g(X[x]), X)$. It is not protected in $f(g(X[x]), x)$ because
 409 of its second occurrence. Protected variables can be eliminated from a term by appropriately
 410 instantiating its meta-variables, while unprotected variables cannot be eliminated. An important
 411 known observation to be justified later is that elementary unification problems for which a free
 412 variable occurs unprotected on one side, and does not occur at all on the other side have no solution.

413 Pattern matching and unification are described by the rewrite rules given at Figure 1. Rules
 414 written in black apply to both matching and unification problems. Rules written in green apply
 415 to matching problems only, while rules in blue apply to unification problems only. Note that the
 416 constant \perp is absorbing, a black rule that will remain implicit.

417 The initial problem to be matched or unified is denoted by P_0 . Rule *Fail-Protect* refers to P_0 .

418 Apart from *Meta-Var*, the set of common black rules treats equations between expressions which
 419 are not pre-redexes. Those equations can be decomposed or fail, depending on the respective root
 420 symbol of the left-hand and right-hand sides. These rules are just the same as those for first-order
 421 unification. The role of *Meta-Var* is to ensure that the arity condition for meta-variables is met by
 422 the substitution that will be obtained if the algorithm succeeds.

423 The two green rules for matching are failure rules. *Fail-Arity* applies when the arity condition
 424 for meta-variables cannot be met, while *Fail-Protect* applies as soon as there is an equation whose
 425 right-hand side contains a free variable that does not occur in the left-hand side or in the initial
 426 problem. The two failure rules for unification require different conditions, in particular because
 427 unification is symmetric while matching is not. *Fail-Arity* treats equations which falsify the arity
 428 condition for meta-variables, while *Fail-Protect* deals with equations which cannot be unified
 429 because the right-hand side has an unprotected variable. Note that u cannot be a pre-redex in that
 430 case.

431 There are three remaining blue rules for unification. When the right-hand side of an equation is
 432 a pre-redex, it is moved to the left by *Swap* if the left-hand side is not a pre-redex, or else by *Flip*
 433 if the left-hand side is lacking more (implicit) arguments than the right-hand side. *Drop* applies to
 434 equations with a pre-redex on the left, in case there is some protected variable in the right-hand
 435 side that must be eliminated, as stipulated by the first condition. The other three conditions, in the
 436 order they are listed, postpone the application of *Drop* until *Fail-Arity*, *Fail-Protect* and *Flip*, in this
 437 order, can no longer apply.

438 In the particular case where $q = \wedge$, $|X| - |\bar{x}| = |Y| - |\bar{y}|$ and $\bar{x} \subseteq \bar{y}$ then *Drop* applies to $X[\bar{x}] = Y[\bar{y}]$
 439 and produces $X[\bar{x}] = Z[\bar{x}] \wedge Y[\bar{y}] = Z[\bar{x}]$ which could be improved in an implementation with a
 440 special instance of *Drop* to produce $Y[\bar{y}] = X[\bar{x}]$ only, as does *Flip*.
 441

442			
443	<i>Dec-Fun</i>	$f(\bar{u}) = f(\bar{v}) \longrightarrow \bigwedge_{i=1}^{ \bar{f} } u_i = v_i$	if $f \in \mathcal{F} \cup \mathcal{X}$
444	<i>Dec-App</i>	$(u\ s) = (v\ t) \longrightarrow u = v \wedge s = t$	if $(u\ s)$ and $(v\ t)$ are not pre-redexes
445	<i>Dec-Abs</i>	$\lambda x.u = \lambda y.v \longrightarrow u\{x \mapsto z\} = v\{y \mapsto z\}$	with z fresh
446	<i>Conflict</i>	$f(\bar{u}) = g(\bar{v}) \longrightarrow \perp$	if $f, g \in \mathcal{F} \cup \mathcal{X} \cup \{\@, \lambda\}$, $f \neq g$
447	<i>Meta-Abs</i>	$X[\bar{x}] = \lambda y.v \longrightarrow X[\bar{x}y] = v$	if $X \in \mathcal{Z}$, $ X > \bar{x} $
448	<i>Fail-Arity</i>	$X[\bar{x}] = u \longrightarrow \perp$	if $ X - \bar{x} > ar(u)$
449	<i>Fail-Protect</i>	$X[\bar{x}] = u \longrightarrow \perp$	if $\exists z \in \mathcal{V}ar(u)$, $z \notin \bar{x} \cup \mathcal{V}ar(P_0)$
450	<i>Fail-Arity</i>	$X[\bar{x}] = f(\bar{u}) \longrightarrow \perp$	if $ X > \bar{x} \wedge f \in \{\@\} \cup \mathcal{F} \cup \bar{x}$
451	<i>Fail-Protect</i>	$X[\bar{x}] = u \longrightarrow \perp$	if $\exists z \in \mathcal{V}ar(u)$, $z \notin \bar{x}$, z unprotected
452			
453	<i>Swap</i>	$u = Y[\bar{y}] \longrightarrow Y[\bar{y}] = u$	if u is not a pre-redex
454	<i>Flip</i>	$X[\bar{x}] = Y[\bar{y}] \longrightarrow Y[\bar{y}] = X[\bar{x}]$	if $ X - \bar{x} > Y - \bar{y} $
455			$\left\{ \begin{array}{l} \bar{y} \notin \bar{x} \cup \mathcal{B}Var(u) \\ u(\Lambda) \notin \mathcal{F} \cup \{\@\, \lambda\} \vee X = \bar{x} \\ \text{unprotected variables of } u \text{ are in } \bar{x} \\ q \neq \Lambda \vee Y - \bar{y} \geq X - \bar{x} \end{array} \right.$
456		$\left\{ \begin{array}{l} \bar{z} = \bar{y} \cap (\bar{x} \cup \mathcal{B}Var(u)) \\ Z \text{ fresh s.t. } Z = Y - \bar{y} + \bar{z} \end{array} \right.$	
457	<i>Drop</i>		
458			
459			
460			
461			
462			
463			
464			
465			
466			
467			
468			
469			
470			
471			
472			
473			
474			
475			
476			
477			
478			
479			
480			
481			
482			
483			
484			
485			
486			
487			
488			
489			
490			

Fig. 1. Matching and unification rules for linear equational problems

Note that the set of rules can be easily transformed into a deterministic algorithm as no two rules apply to the same equation, except *Meta-Abs* and each of the failure rules but *Fail-Arity*.

Before to prove properties of these matching and unification rules, we show below examples of use of the unification rules that are useful for the reader's understanding:

Example 4.5. Let's illustrate some rules, using $|X| = 1$, $|Y| = 0$, $|X'| = 3$, $|Y'| = 2$ and $|Z| = 2$.

$$\begin{aligned}
 f(\lambda y.f(Y)) = f(X) &\xrightarrow{\text{Dec-Fun}} \lambda y.f(Y) = X \xrightarrow{\text{Swap}} X = \lambda y.f(Y) \xrightarrow{\text{Meta-Var}} X[y] = f(Y) \\
 f(Y') = f(\lambda y.f(Y)) &\xrightarrow{\text{Dec-Fun}} Y' = \lambda y.f(Y) \xrightarrow{\text{Meta-Var}} Y'[y] = f(Y) \xrightarrow{\text{Fail-Arity}} \perp \\
 X' = \lambda y.Y'[y] &\xrightarrow{\text{Meta-Var}} X'[y] = Y'[y] \xrightarrow{\text{Flip}} Y'[y] = X'[y] \\
 Y'[z] = \lambda x.X'[y, z] &\xrightarrow{\text{Dec-var}} Y'[z, x] = X'[y, z] \xrightarrow{\text{Drop}} Y'[z, x] = Z[z] \wedge X'[y, z] = Z[z]
 \end{aligned}$$

Drop applies here to an elementary equation in which there are extra-variables on both sides, eliminating, perhaps surprisingly, both problems at once: the two generated equations have a pre-redex on the left-hand side which contains all free variables occurring on the other side.

We now show examples of matching and unification problems that will be useful later when computing the critical pairs of the theory of global states. We won't do all computations needed later on, only a few interesting ones originating from the first or second versions of that theory:

Example 4.6. We start with two matching problems, matching first the term $lk(\lambda w.ud(w, X'[w]))$ with the left-hand side $lk(\lambda w.ud(w, X[w]))$ of rule (lu) of the second set:

$$\begin{aligned}
 & lk(\lambda w.ud(w, X[w])) = lk(\lambda w.ud(w, X'[w])) \xrightarrow{Dec-Fun} \lambda w.ud(w, X[w]) = \lambda w.ud(w, X'[w]) \\
 & \xrightarrow{Dec-Abs} ud(w, X[w]) = ud(w, X'[w]) \xrightarrow{Dec-Fun} w = w \wedge X[w] = X'[w] \xrightarrow{Dec-Fun} X[w] = X'[w]
 \end{aligned}$$

Now, the second, matching the term $lk(\lambda v.lk(Y'[v]))$ against the left-hand side $lk(\lambda v.lk(Y'[v]))$ of rule (ll) of the second set:

$$\begin{aligned}
 lk(\lambda v.lk(Y'[v])) = lk(\lambda v.lk(Y'[v])) & \xrightarrow{Dec-Fun} \lambda v.lk(Y'[v]) = \lambda v.lk(Y'[v]) \xrightarrow{Dec-Abs} \\
 lk(Y'[v]) = lk(Y'[v]) & \xrightarrow{Dec-Fun} Y[v] = Y'[v]
 \end{aligned}$$

It then follows that $lk(\lambda v.lk(Y'[v])) \xrightarrow{ll} lk(\lambda v.Y'[v, v])$.

We go on with unification problems, first of the left-hand sides of rules (ll) and (l) from the first set:

$$\begin{aligned}
 lk(\lambda w.lk(\lambda v.Y[w, v])) = lk(\lambda w.U) & \xrightarrow{Dec-Fun} \lambda w.lk(\lambda v.Y[w, v]) = \lambda w.U \xrightarrow{Dec-Abs} \\
 lk(\lambda v.Y[w, v]) = U & \xrightarrow{Swap} U = lk(\lambda v.Y[w, v]) \xrightarrow{Drop} U = lk(\lambda v.Z[v]) \wedge Y[w, v] = Z[v]
 \end{aligned}$$

(with Z fresh of arity 1)

Now, we consider the unification of the left-hand sides of rules (lu) and (l) from the same set:

$$\begin{aligned}
 lk(\lambda v.ud(v, X[v])) = lk(\lambda v.U) & \xrightarrow{Dec-Fun} \lambda v.ud(v, X[v]) = \lambda v.U \xrightarrow{Dec-Abs} \\
 ud(v, X[v]) = U & \xrightarrow{Swap} U = ud(v, X[v]) \xrightarrow{Fail-Protect} \perp
 \end{aligned}$$

since v occurs unprotected as first argument of ud in $ud(v, X[v])$, making unification fail.

Finally comes unification of the left-hand sides of rule (l) with a subterm of the left-hand side of rule (ul) (still from the same set):

$$lk(\lambda v.U) = lk(\lambda v.X[v]) \xrightarrow{Dec-Fun} \lambda v.U = \lambda v.X[v] \xrightarrow{Dec-Abs} U = X[v] \xrightarrow{Swap} X[v] = U$$

We can now carry out the same computations using the second set of rules. We get first:

$$\begin{aligned}
 lk(\lambda w.lk(Y[w])) = lk(\lambda w.U) & \xrightarrow{Dec-Fun} \lambda w.lk(Y[w]) = \lambda w.U \xrightarrow{Dec-Abs} lk(Y[w]) = U \xrightarrow{Swap} U = lk(Y[w]) \\
 & \xrightarrow{Drop} U = lk(Z) \wedge Y[w] = Z \quad (\text{with } Z \text{ fresh of arity 1})
 \end{aligned}$$

The second computation is exactly the same. We move to the third:

$$lk(\lambda v.U) = lk(X) \xrightarrow{Dec-Fun} \lambda v.U = X \xrightarrow{Swap} X = \lambda v.U \xrightarrow{Meta-Var} X[v] = U$$

We observe that the computations from the second set of rules are identical in the first two cases, but slightly different in the third case. On the contrary, the obtained solved form is the same in the third case, but slightly different in the first case for which the second one can be deduced from the first by an $M\eta$ -step.

We now go on studying the matching/unification rules. First, we verify that the rules operate on equational problems (linearity will be considered later):

LEMMA 4.7. *Assume that an equational problem P rewrites to P' by using one of the matching/unification rules. Then, P' is an equational problem.*

PROOF. All rules preserve the property that pre-redexes are never applied. \square

540 The following sequence of properties shows that all rules are sound, that is preserve the solutions
541 of equational problems.

542 We start with the three rules checking arities. We have seen that arities are non-decreasing under
543 substitutions, and are even sometimes preserved. This is the basis for the soundness of the rules:

544 LEMMA 4.8. *Assume that $|X| > |\bar{x}|$. Then the elementary unification problems $X[\bar{x}] = \lambda y.v$ and
545 $X[\bar{x}y] = v$ have the same set of solutions.*

546 PROOF. Without loss of generality, we assume that $y \notin \bar{x}$ and restrict our attention to solutions
547 σ such that $y \notin \text{Dom}(\sigma)$. Since $|X| > |\bar{x}|$, a substitution for X must be of the form $\{X \mapsto \lambda \bar{x}z.u\}$.
548 Such a substitution σ is a solution of $X[\bar{x}] = \lambda y.v$ iff $(\lambda z.u)\{\bar{x} \mapsto \bar{x}\} = \lambda z.u = (\lambda y.v)\sigma = \lambda y.v\sigma$,
549 which holds true iff $u\{z \mapsto y\} = v\sigma$, which in turn holds true iff σ is a solution of $X[\bar{x}y] = v$. \square

550 LEMMA 4.9. *The elementary matching problem $X[\bar{x}] = u$ has no solution if $\text{ar}(X[\bar{x}]) > \text{ar}(u)$.*

551 PROOF. By non-decreasingness, $\text{ar}(X[\bar{x}]\sigma) \geq \text{ar}(X[\bar{x}]) > \text{ar}(u)$. \square

552 LEMMA 4.10. *Assume that $|X| > |\bar{x}|$ and $f \in \mathcal{F} \cup \{\text{@}\} \cup \mathcal{X}$. Then, the elementary unification
553 problem $X[\bar{x}] = f(\bar{u})$ has no solution.*

554 PROOF. Again, $\text{ar}(X[\bar{x}]\sigma) \geq \text{ar}(X[\bar{x}]) > 0 = \text{ar}(f(\bar{u}\sigma)) = \text{ar}(f(\bar{u})\sigma)$, since $x\sigma = x$ when $f \in \mathcal{X}$. \square

555 We now move to the case where extra-variables occur in right-hand sides, whether protected or
556 not, starting with the cases of unprotected variables:

557 LEMMA 4.11. *Let u be a term containing a variable $z \notin \bar{x} \cup \mathcal{V}\text{ar}(P)$. Then, the elementary matching
558 problem $X[\bar{x}] = u$ has no solution in common with \mathcal{P} .*

559 PROOF. Assume γ is a common solution for $X[\bar{x}] = u$ and \mathcal{P} . By definition of a solution of \mathcal{P} ,
560 $\gamma(X) = \lambda \bar{x}.v$ with $\mathcal{V}\text{ar}(v) \subseteq \bar{x} \cup \mathcal{V}\text{ar}(\mathcal{P})$. By assumption on z , $X[\bar{x}]\gamma$ and u have different sets of
561 free variables, hence $X[\bar{x}]\gamma \neq_\alpha u$, hence contradicting our assumptions. \square

562 LEMMA 4.12. *Let u be a term containing an unprotected variable $z \notin \bar{x}$. Then, the elementary
563 unification problem $X[\bar{x}] = u$ has no solution.*

564 PROOF. Assume γ is a solution for $X[\bar{x}] = u$. By definition of a solution, $\gamma(X) = \lambda \bar{x}.v$ with
565 $\mathcal{V}\text{ar}(v) \subseteq \bar{x}$, hence $z \notin \mathcal{V}\text{ar}(X[\bar{x}]\gamma)$. By definition again, $z \notin \text{Dom}(\gamma)$, hence $z\gamma = z$, and therefore
566 $z \in \mathcal{V}\text{ar}(u\gamma)$. Since $X[\bar{x}]\gamma$ and $u\gamma$ have different sets of free variables, no α -renaming can make
567 them equal, hence contradicting our assumption. \square

568 LEMMA 4.13. *Let E be the elementary unification problem $X[\bar{x}] = u[Y[\bar{y}]]_q$ and P the unification
569 problem $X[\bar{x}] = u[Z[\bar{z}]]_q \wedge Y[\bar{y}] = Z[\bar{z}]$, where $\bar{z} = \bar{y} \cap (\bar{x} \cup \mathcal{B}\mathcal{V}\text{ar}(u))$ and Z is a fresh variable of
570 arity $|Y| - |\bar{y}| + |\bar{z}|$. Then, γ is a solution of P iff $\gamma_{\bar{z}}$ is a solution of E .*

571 PROOF. Let \bar{x}' (resp., \bar{y}' , \bar{z}') be a vector of pairwise distinct fresh variables of length $|X| - |\bar{x}|$
572 (resp., $|Y| - |\bar{y}|$, $|Z| - |\bar{z}|$), and $\gamma = \{X \mapsto \lambda \bar{x}x'.w, Y \mapsto \lambda \bar{y}y'.w', Z \mapsto \lambda \bar{z}z'.w''\}$ be a solution of the
573 generated problem. Remark that \bar{y}' and \bar{z}' have the same length so that we could actually identify
574 them. By definition of a substitution, we get $\lambda \bar{x}'w = u\gamma[\lambda \bar{z}'w'']_q$ and $\lambda \bar{y}'w' = \lambda \bar{z}'w''$, hence
575 $\lambda \bar{x}'w = u\gamma[\lambda \bar{y}'w']_q$, showing that $\gamma_{\bar{z}}$ is a solution to the original problem.

576 Conversely, let $\gamma = \{X \mapsto \lambda \bar{x}x'.w, Y \mapsto \lambda \bar{y}y'.w'\}$ be a solution of the original problem. Using
577 the previous remark, we can tentatively extend γ as γ' by letting $\gamma'(Z) = \lambda \bar{z}z'.w'$. By definition of
578 substitutions, $\lambda \bar{x}'w = u\gamma'[\lambda \bar{y}'w']_q$. It follows that both sides of the equations have the same set of
579 free variables, hence $\mathcal{V}\text{ar}(w') \subseteq \bar{y}' \cup \mathcal{B}\mathcal{V}\text{ar}(u) \cup (\mathcal{V}\text{ar}(w) \setminus \bar{x}') \subseteq \bar{y}' \cup \mathcal{B}\mathcal{V}\text{ar}(u) \cup \bar{x} \cup \mathcal{V}\text{ar}(E) \subseteq$

589 $\overline{y'} \cup \overline{z} \cup \mathcal{V}ar(E)$. Hence $\mathcal{V}ar(\lambda\overline{z}y'.w') \subseteq \mathcal{V}ar(E) = \mathcal{V}ar(P)$, and y' satisfies the requirements to
 590 be a candidate solution of P . Showing that it satisfies P is routine. \square

591

We can now conclude:

592

593

594

595

LEMMA 4.14. *The matching/unification rules are terminating and preserve solutions of matching and unification problems.*

596

597

598

599

600

601

602

603

604

605

PROOF. Termination of the matching rules is clear. For unification, we interpret an equational problem by the multiset of interpretations of its elementary equations, so that it is enough to show that the interpretation of each elementary equation generated by a unification rule is strictly less than the interpretation of its left-hand side. An elementary equation $u = v$ is interpreted by the quadruple $\langle m, n, p, q \rangle$, where m is the size of the equation from which all pre-redexes have been removed, $n = 1$ if u is not a pre-redex otherwise 0, and $p = 1$ if $ar(u) > ar(v)$ otherwise 0, and q is the number of variables occurrences in v . It is easy to see that all rules but the last three generate elementary equations whose interpretation's first component has decreased strictly. Now, *Swap*, *Flip*, *Drop* decrease respectively their second, third, fourth component without changing their previous ones. In the case of *Drop*, this follows from the easy property that $\overline{z} \subseteq \overline{y}$.

606

607

608

609

610

Preservation of solutions follows from: for the first four rules, the fact that they apply above the fringe since they can't apply to an equation resulting from any other rule application; Lemma 4.8 for *Meta-Var*; Lemma 4.9 for *Fail-Arity*; Lemma 4.11 for *Fail-Protect*; Lemma 4.10 for *Fail-Arity*; Lemma 4.12 for *Fail-Protect*; commutativity of equality for *Swap* and *Flip*; and Lemma 4.13 for *Drop* (preservation is relative here to the variables of the initial problem). \square

611

612

We now give the characterization of equational problems in normal form for those rules:

613

Definition 4.15 (Solved forms).

614

615

616

617

618

- (1) An equation $u = v$ is in *arity solved form* if u is a pre-redex such that $ar(u) \leq ar(v)$;
- (2) An equational problem is in *solved form* if it is either the constant \perp , or a conjunction $\bigwedge X_i[\overline{x}_i] = v_i$ of equations in *arity solved form* such that for all i , $\mathcal{V}ar(v_i) \subseteq \overline{x}_i$ and for all i, j , $X_i \notin \mathcal{M}\mathcal{V}ar(v_j)$ and for all $i \neq j$, $X_i \neq X_j$.

619

LEMMA 4.16.

620

621

622

623

- (1) *An equation is in arity solved form iff it is irreducible by all rules but Fail-Protect, Fail-Protect, Drop and Meta-Abs;*
- (2) *Drop and Meta-Abs preserve arity-solved forms.*

624

625

626

627

628

629

PROOF. (1) The only if case being clear, let us assume that no rule other than *Fail-Protect*, *Fail-Protect*, *Drop* and *Meta-Abs* can apply $u = v$. Necessarily $u = X[\overline{x}]$ otherwise one of the *Dec* rules, *Conflict* or *Swap* rules would apply. Assuming $|\overline{x}| < |X|$, then v must be some pre-redex $Y[\overline{y}]$ otherwise *Meta-Abs*, *Fail-Arity* or *Fail-Arity* would apply. Because *Flip* doesn't apply by assumption, we conclude that $|X| - |\overline{x}| \leq |Y| - |\overline{y}|$, hence $u = v$ is in arity solved form.

630

631

632

633

634

(2) The application conditions of *Drop* imply that the *Dec* rules, *Conflict*, *Meta-Var*, *Swap* and *Flip* don't apply. From (1), it follows that *Drop* applies only on equations in arity solved form.

(3) The case of *Meta-Abs* follows from the definition of arity. Consider now *Drop*. If $q = \wedge$, then $|X| - |\overline{x}| \leq |Y| - |\overline{y}| = |Z| - |\overline{z}|$ and the first generated equation is in arity solved form. Otherwise $|X| = |\overline{x}|$. The first generated equation is in arity solved form, as was the input equation. The second generated equation is in arity solved form because $|Z| - |\overline{z}| = |Y| - |\overline{y}|$. \square

635

636

637

LEMMA 4.17. *All rules preserve the following two properties of an unification problem P :*

- 638 (1) For all equation not in arity solved form $u = v \in P$, meta-variables in $\mathcal{MVar}(u) \cup \mathcal{MVar}(v)$
 639 are linear in P .
 640 (2) For all equation in arity solved form $X[\bar{x}] = v \in P$, X is linear in P and all occurrences of a
 641 non-linear meta-variable Z in v are exclusively applied to variables in $\mathcal{Var}(v) \cup \bar{x}$.

642 **PROOF.** All rules but *Meta-Abs*, *Fail-protect* and *Drop* preserve linearity of elementary equations,
 643 hence both properties (1) and (2) are preserved in that case. The case of *Meta-Abs* and *Fail-Protect*
 644 is clear. We are left with *Drop* which operates on equations in arity solved form. From (2), non-
 645 linear meta-variables in right-hand sides are applied to locally bound or left-hand side variables,
 646 which prevents the application of *Drop* in that case. On the other, any application of *Drop* to linear
 647 meta-variable produces equations that satisfy the property since $\bar{z} \subseteq \bar{y}$ and $\bar{z} \subseteq \bar{x} \cup \mathcal{BVar}(u)$. \square
 648

649 **LEMMA 4.18.** *Let P be a linear (matching or unification) equational problem: its normal form is in*
 650 *solved form.*

651 **PROOF.** Let Q be the normal form of P . By Lemma 4.16 (1), all equations in Q are in arity-solved
 652 form. By Lemma 4.17 (2), left-hand sides pre-redexes are linear in Q . We are left proving that for
 653 all $X[\bar{x}] = v \in Q$, $\mathcal{Var}(v) \subseteq \bar{x}$. Assume it is not the case, and let $y \in \mathcal{Var}(v) \setminus \bar{x}$. Either y occurs
 654 protected and *Drop* applies, or else *Fail-Protect* applies. \square
 655

656 **LEMMA 4.19.** *Solved forms of equational problems are computed in linear time.*

657 **PROOF.** First, note that the matching and unification rules check a fixed number of symbols
 658 belonging to the head of the left-hand and right-hand side of each equation belonging to a given
 659 equational problem in turn. It is therefore enough to show that the total number of matching or
 660 unification steps is linear in the size of the problem. Finally, because meta-variables appear linearly
 661 in a given equational problem, it is enough to consider every elementary equation separately.
 662

663 The set of rules common to unification and matching applies to an elementary equation $u = v$ a
 664 number of times bound by $|\mathcal{FPos}(u)|$, and yields a number of elementary equations whose whole
 665 size is bound by $|u| + |v|$. Failure rules may apply only once. This concludes the case of matching,
 666 we continue with the remaining unification rules. *Swap* and *Flip* may apply only once to a given
 667 equation, and leave the size of the problem invariant. Finally, the conditions for applying *Drop*
 668 ensure that no other rule will ever apply after any sequence of *Drops*. Further the total number
 669 of applications of *Drop* to a given equation is bound by the number of protected variables to be
 670 eliminated, hence by its size. This shows that the whole unification process is linear. \square

671 We are left extracting most-general unifiers from equational problems in solved form:

672 **LEMMA 4.20.** *A solved form $P = \wedge_i X_i[\bar{x}_i] = v_i$ has a unique (up to renaming of bound variables)*
 673 *most general unifier $\text{mgs}(P) = \{X_i \mapsto \lambda \bar{x}_i \lambda \bar{z}. (v_i \bar{z})\}_i$, that is, every ground solution of P is a ground*
 674 *instance of σ .*
 675

676 **PROOF.** First, $X_i[\bar{x}_i]\sigma = \lambda \bar{z}. (v_i \bar{z})$. On the other hand, by definition of a solved form, $v_i\sigma = v_i$.
 677 Hence $X_i[\bar{x}_i]\sigma =_{\alpha \cup M\eta} v_i\sigma$, hence σ is a unifier of the solved form.

678 Instantiating the equation $v_i =_{M\eta} X_i[\bar{x}_i]\sigma$ with an arbitrary solution γ of P , we get $v_i\gamma =$
 679 $X_i[\bar{x}_i]\sigma\gamma$ ($M\eta$ -steps disappear, as stressed in subsection 3.2). Using now the fact that γ is a solution
 680 yields $X_i[\bar{x}_i]\gamma = X_i[\bar{x}_i]\sigma\gamma$, showing that γ is an instance of σ by itself (as in the first-order case). \square
 681

682 **Example 4.21.** Consider the two solved forms obtained at example 4.6: $U = lk(\lambda v. Z[v]) \wedge$
 683 $Y[w, v] = Z[v]$ and $U = lk(Z) \wedge Y[w] = Z$, where $|Z| = 1$ for both cases. The most general solution
 684 is $\{U \mapsto lk(\lambda v. Z[v]), Y \mapsto \lambda w v. Z[v]\}$ for the first, and $\{U \mapsto lk(Z), Y \mapsto \lambda w. Z\}$ for the second.
 685 Consider now the matching solved form obtained at example 4.6: $X[w] = X'[w]$. The matching
 686

substitution obtained is $\{X \mapsto \lambda w.X'[w]\}$. It would of course be possible to extract the better mgs $\{X \mapsto X'\}$ to the price of some more technicalities.

Pattern matching and unification of linear patterns is therefore quite easy: first, reduce the initial pattern matching problem $L = u$, or unification problem $L|_p = L'|_{p'}$, to a solved form. Then extract the matching substitution or the most general unifier from the solved form. Therefore:

THEOREM 4.22. *The matching problem results in a single matching substitution when it succeeds, computable in linear time. The unification problem results in a single (up to α) most general solution when it succeeds, computable in linear time.*

Note that there is no mention of types in this algorithm. A natural question is whether the most general solution is typed when two linear patterns get unified, and whether it coincides with the one obtained when unifying dependently typed linear patterns. This question will be addressed in a forthcoming paper, in which the linearity restriction on patterns is removed.

5 LOCAL PEAKS IN REWRITE THEORIES

Rewrite theories have two kinds of local ancestor peaks, *homogeneous* ones, between functional or higher-order reductions, and *heterogeneous* ones, which mix both kinds of reductions. We analyze here which local ancestor peaks enjoy *decreasing diagrams for free*, and which do not. Some results in this section will be reused in forthcoming papers, those that do not rely on the left-linearity assumption for the rewrite rules, nor on orthogonal functional reductions.

5.1 Decreasing diagrams for free

A key property of plain first-order rewriting is that there are three possible kinds of local peaks depending on the respective positions of the rewrites that define them. This property generalizes trivially to higher-order rewrites with our definition of set of positions for patterns:

LEMMA 5.1. *Given terms s, t such that $s \xleftarrow[p]{i:L \rightarrow R} u \xrightarrow[q]{j:G \rightarrow D} t$, then, there are three possibilities: (i) $p \# q$ (disjoint peak case); (ii) $q \geq_p p \cdot F_L$ or $p \geq_p q \cdot F_L$ (ancestor peak case); and (iii) $p = q \cdot o$ with $o \in \mathcal{FPos}(L)$ or $q = p \cdot o$ with $o \in \mathcal{FPos}(G)$ (overlapping peak case).*

In the case of plain rewriting, two non-overlapping rewrite steps issuing from a same term commute, a major component of any confluence proof. When the steps occur at disjoint positions, this property, which holds for any monotonic relation, remains true for rewriting modulo a theory, hence all disjoint peaks have decreasing diagrams for free. This is not the case, however, when the steps occur at positions whose one is an ancestor of the other, because the modulo part of the above rewrite may interact with the rewrite below. Our definition of higher-order rewriting, however, enjoys a similar property, because the fringe of a rewrite step protects positions below it.

In the coming lemma, “LAP” stands for *linear ancestor peak*, and “a” for *above*, the β -step being above a higher-order step. It applies to any higher-order rule, left-linear or not.

LEMMA 5.2 (LAP β A). *Let u be a term, $p, q \in \mathcal{Pos}(u)$ such that $q \geq_p p \cdot F_\beta$ and $s \xleftarrow[p]{i:L \rightarrow R} u \xrightarrow[q]{j:G \rightarrow D} t$. Then $s \xrightarrow[\beta]{j} \xleftarrow[p]{i} t$ for some set Q of parallel positions of s such that $Q \geq_p p$.*

PROOF. By assumption, $u_p = s_p = t_p$, $u|_p = (\lambda x.M N)$, and $s|_p = M\sigma$, where $\sigma = \{x \mapsto N\}$. There are two cases:

The case where $q = p \cdot 2 \cdot q'$ and $t|_p = (\lambda x.M P)$ with $N \xrightarrow{j} P$. This requires several j -steps at the parallel positions of x in M . Then $s = u[M\sigma]_p \xrightarrow{j} u[M\{x \mapsto P\}]_p \xleftarrow{\beta} u[(\lambda x.M P)]_p = t$.

Otherwise, $q = p \cdot 1^2 \cdot q'$, that is, $M|_{q'} = u|_q \xrightarrow{j} t|_q$. Then, $s = u[M\sigma[u|_q\sigma]_{q'}]_p$. By Lemma 3.12, $u|_q\sigma \xrightarrow{j} t|_q\sigma$, hence, by Lemma 3.11, $s \xrightarrow{j} u[M\sigma[t|_q\sigma]_{q'}]_p$. On the other hand, $t|_p = (\lambda x.P N)$, where $P = M[t|_q]_{q'}$, therefore $t = u[(\lambda x.P N)]_p \xrightarrow{\beta} u[P\sigma]_p = u[M\sigma[t|_q\sigma]_{q'}]_p$. We are done. \square

The case of a local peak $s \xleftarrow{i} u \xrightarrow{j} t$, where the higher-order step with $i : L \rightarrow R$ applies above another step belonging to $\mathcal{R} \cup \beta$, a situation called (LAPRa), is shown at Figure 3 (left). (LAPRa) does not apply to non-left-linear rules. Its proof requires an important preliminary result:

LEMMA 5.3 (PRESERVATION). *Let $u \xrightarrow{i:L \rightarrow R} v$ with L linear and $q \in \text{Pos}(u)$ such that $q \geq_{\mathcal{P}} p \cdot F_L$. Then $\underline{u}_q = u[Z[\bar{z}]]_q \xrightarrow{i} w$ for some w , where \bar{z} is the list of variables bound above q in u , Z fresh, $|Z| = |\bar{z}|$, and $v = w\bar{u}^q = w\{Z \mapsto \lambda\bar{z}.u|_q\}$.*

PROOF. By definition of splitting, $u = t\tau$, where $t = \underline{u}_q = u[Z[\bar{z}]]_q$ and $\tau = \bar{u}^q = \{Z \mapsto \lambda\bar{z}.u|_q\}$.

Since $q \geq_{\mathcal{P}} p \cdot F_L$, then $q = p \cdot o \cdot q'$, where $o \in F_L$ is the position of a pre-redex in L . Hence $L|_o = X[\bar{x}]$ for some meta-variable X and variables \bar{x} bound above o in L .

By definition of higher-order rewriting, $u|_p = L\gamma$ for some substitution γ . By definition of a substitution, $\gamma(X) = \lambda\bar{x}.M$, and by the previous property, $X[\bar{x}]\gamma = M = u|_{p \cdot o}$, hence $M|_{q'} = u|_q$. As a consequence, $u|_{p \cdot o}[Z[\bar{z}]]_{q'} = M[Z[\bar{z}]]_{q'}$.

Let now θ be the substitution identical to γ except for the meta-variable X for which $\theta(X) = \lambda\bar{x}.M[Z[\bar{z}]]_{q'}$. We have $\theta(X)\tau = \lambda\bar{x}.M[u|_q]_{q'} = \lambda\bar{x}.M[M|_{q'}]_{q'} = \lambda\bar{x}.M = \gamma(X)$, hence $\gamma = \theta\tau$.

Since L is linear, there is a single pre-redex containing the meta-variable X . As a consequence, $L\theta = u|_p[X[\bar{x}]\theta]_o = u|_p[(M[Z[\bar{z}]]_{q'})_o]_p = u|_p[u|_{p \cdot o}[Z[\bar{z}]]_{q'}]_o = u|_p[Z[\bar{z}]]_{o \cdot q'}$, and therefore $\underline{u}_q = u[L\theta]_p$.

By definition of higher-order rewriting, $L\theta \xrightarrow{i} R\theta$, and by monotonicity, $\underline{u}_q \xrightarrow{i} u[R\theta]_p = w$.

By definition of higher-order rewriting again, $v = u[R\gamma]_p = u[R\theta\tau]_p = (u[R\theta]_p)\tau = w\bar{u}^q$. \square

As already said, (LAPRa) requires the linearity assumption.

LEMMA 5.4 (LAPRa). *Let \mathcal{R} be a left-linear rewrite system, $i : L \rightarrow R \in \mathcal{R}$, $j \in \mathcal{R} \cup \beta$, u be a term, and $p, q \in \text{Pos}(u)$ such that $q \geq_{\mathcal{P}} p \cdot F_L$ and $s \xleftarrow{i} u \xrightarrow{j} t$. Then, $s \xrightarrow{j} \xrightarrow{\geq_{\mathcal{P}} p} \xleftarrow{i} t$.*

PROOF. Splitting u at q yields $u = v\sigma$, where $v = \underline{u}_q = u[Z[\bar{z}]]_q$ and $\sigma = \bar{u}^q = \{Z \mapsto \lambda\bar{z}.u|_q\}$ is preserving since $u|_q$ cannot be an abstraction by definition of a pattern and is normal as a subterm of $u|_p$. By assumption, $u|_q \xrightarrow{j} t|_q$, and by monotonicity, $\sigma(Z) = \lambda\bar{z}.u|_q \xrightarrow{j} \lambda\bar{z}.t|_q$. Let τ be σ with the exception $\tau(Z) = \lambda\bar{z}.t|_q$. Then $\sigma \xrightarrow{j} \tau$ and by Lemma 3.11, $u = v\sigma \xrightarrow{j} v\tau = t$. By Lemma 5.3, $v \xrightarrow{i} w$ for some w such that $s = w\sigma$. By Lemma 3.13, $u = v\sigma \xrightarrow{j} w\tau$. The result follows. \square

This ancestor peak property is more complex than for first-order computations since we need the assumption that the rewrite rules are left-linear, which is of course true of all first-order rewrite rules, and is true as well of all rules of the theory of global states.

Note that whenever X occurs self-nested in a right-hand-side, $R[X[\bar{u}]]_p$, such that $\bar{u}_i^q = \lambda \bar{z}.X[\bar{t}]$ for some i , this right-hand-side can be replaced with $R[(\lambda x.X[\bar{v}] \lambda \bar{z}.X[\bar{t}])]_p$ with $v_j = u_j$ for $i \neq j$ and $v_i = u_i[(x \bar{z})]_q$. For instance, instead of using the rule $f(\lambda x.X[x]) \rightarrow X[g(\lambda z.X[h(z)])]$ one may use $f(\lambda x.X[x]) \rightarrow (\lambda x.X[g(\lambda z.(x z))] \lambda z.X[h(z)])$. In case the rule has critical pairs, they will most presumably require to be joined with β -steps which make this transformation of little interest.

5.2 Critical peaks

We can now define critical peaks:

Definition 5.5. Let $i : L \rightarrow R$ and $j : G \rightarrow D$ be two rules in \mathcal{R} and $o \in \mathcal{FPos}(L)$ such that the equation $L|_o = G$ has a most general solution σ . Then, the peak $R\sigma \xleftarrow{i} L\sigma \xrightarrow{o} L\sigma[D\sigma]_o$ is called a *critical peak* of j onto i at position o . Its associated *critical pair* is $\langle R\sigma, L\sigma[D\sigma]_o \rangle$.

This definition makes sense: since $o \in \mathcal{FPos}(L)$, then $o <_{\mathcal{P}} F_L$, and therefore, $o \in \mathcal{FPos}(L\sigma)$. Using standard techniques, we then get the analog of Nipkow's critical pair lemma developed for the case of simply typed higher-order rewrite rules:

LEMMA 5.6 (CRITICAL PAIR LEMMA). Assume $s \xleftarrow{i} u \xrightarrow{q} t$ is an overlapping peak of $j : G \rightarrow D$ onto $i : L \rightarrow R$ at position $o \in \mathcal{FPos}(L)$ such that $p = q \cdot o$. Then, there is a critical peak $s' \xleftarrow{i} u' \xrightarrow{o} t'$ and a substitution θ such that $u'\theta = u|_p$, $s'\theta = s|_p$ and $t'\theta = t|_p$.

Thanks to our definition of higher-order rewriting, the proof is similar to the first-order case:

PROOF. By definition of higher-order rewriting, there exists some substitution γ such that $L\gamma = u|_p$, $G\gamma = u|_q$, $s|_p = R\gamma$ and $t|_q = D\gamma$. Since $o \in \mathcal{FPos}(L)$, then $o <_{\mathcal{P}} F_L$, and since $L\gamma = u|_p$, then $u|_p[\]_o = L\gamma[\]_o$, hence $t|_p = u|_p[t|_q]_o = (L\gamma[D\gamma]_o) = L[D]_o\gamma$.

Since $o \in \mathcal{FPos}(L)$, $(L\gamma)|_o = L|_o\gamma$, hence $L|_o\gamma = G\gamma$. Therefore, γ is a solution of the equation $L|_o = G$. Let σ be its most general unifier. Then, there exist a substitution θ such that $\sigma\theta = \gamma$. Hence $u[R\sigma\theta]_p = s$ and $u[L[D]_o\sigma\theta] = t$. Since $s' = R\sigma$ and $t' = L\sigma[D\sigma]_o$, we get the result. \square

5.3 Orthogonal functional reductions

The confluence proof will not be based on using β -rewrites, nor parallel β -rewrites, but, as in Tait's confluence proof of the lambda-calculus, orthogonal β -rewrites (called parallel reductions in [2]). Our definition is essentially Tait's, but makes the rewriting positions explicit.

To this end, we first define the product of sets of positions:

Definition 5.7. Given a set of parallel positions P and a family Q of sets of positions strictly below Λ indexed by P , we define the *orthogonal product* $P \otimes Q$ as the set $P \uplus \bigcup_{p \in P} p \cdot Q_p$.

LEMMA 5.8. Given a set of positions O there exist a unique set of parallel positions $P \subseteq O$ and family $Q < \Lambda$ of sets of positions such that $O = P \otimes Q$.

PROOF. $P = \{p \in S \mid \forall q \in S, p \not\prec q\}$ and $Q_p = \{q > \Lambda \mid p \cdot q \in S\}$ \square

P is called the *parallel* part of O , written \underline{O} , while $\bigcup_{p \in P} p \cdot Q_p$ is called the *residual* part of O , written \bar{O} . Note that $O = \underline{O} \uplus \bar{O}$, $\bar{O} >_{\mathcal{P}} \underline{O}$ and that whenever $O \neq \emptyset$, $\underline{O} \neq \emptyset$ and $\bar{O} \subset O$.

Definition 5.9 (Orthogonal reductions). Orthogonal rewriting, written $u \xrightarrow[\beta]{O} v$, is the smallest reflexive relation on terms such that $u \xrightarrow[\beta]{\bar{O}} s$ and $s \xrightarrow[\beta]{O} v$ imply $u \xrightarrow[\beta]{O} v$.

834 The alternative choice of rewriting first at positions in \underline{O} , instead of first in \overline{O} , would yield a
835 more complex calculation for O , explaining our definition.

836 Note that orthogonal rewriting contains parallel rewriting. Furthermore it is easy to show that
837 our definition of orthogonal reduction coincides with Tait's parallel rewriting. This follows from
838 the property: $t \xrightarrow[\beta]{P} u \xrightarrow[\beta]{Q} v$ implies $t \xrightarrow[\beta]{P \cup Q} v$ if $Q \not\leq_{\mathcal{P}} P$, which is easily proved by induction on P .
839

840 We shall need several well-known properties of (Tait's) orthogonal β -reductions: monotonicity,
841 stability, commutation with any other monotonic rewrite relation, and strong confluence. Besides
842 the following properties will be needed for the coming analysis of orthogonal ancestor peaks.
843

844 LEMMA 5.10. Assume $q \in \mathcal{P}os(u)$ and $O = P \uplus Q \uplus R \subseteq \mathcal{P}os(u)$ with $P \# q$, $Q \geq_{\mathcal{P}} q$ and $R <_{\mathcal{P}} q$.

845 Then $u \xrightarrow[\beta]{O} v$ iff $u \xrightarrow[\beta]{P} \xrightarrow[\beta]{Q} \xrightarrow[\beta]{R} v$.
846

847 PROOF. It suffices to notice that $\xrightarrow[\beta]{P}$ preserves Q and R and $\xrightarrow[\beta]{Q}$ preserves R . \square
848
849

850 LEMMA 5.11. $(\lambda x.M N) \xrightarrow[\beta]{O} t$ with $\Lambda \in O$ if and only if $O = \{\Lambda\} \uplus 1 \cdot P \uplus 2 \cdot Q$, $M \xrightarrow[\beta]{P} M'$,

851 $N \xrightarrow[\beta]{Q} N'$ and $t = M'\{x \mapsto N'\}$.
852
853
854

855 5.4 Orthogonal decreasing diagrams for free

856 We investigate here the linear ancestor peak properties of orthogonal β -reductions. Unlike the
857 "above case", the "below case" listed first follows easily from Lemma 5.2 (LAP β a).
858

859 LEMMA 5.12 (LAPOB). Let $s \xleftarrow[\mathcal{R}ll]{q} u \xrightarrow[\beta]{O} t$ for some set $O >_{\mathcal{P}} q$. Then, $s \xrightarrow[\beta]{\geq_{\mathcal{P}} q} r \xleftarrow[\mathcal{R}ll]{q} t$.

860 Besides $u|_q = L\gamma$ and $s = u[R\gamma]_q$ such that $\gamma \xrightarrow[\beta]{} \sigma$, $t = u[L\sigma]_q$ and $r = u[R\sigma]$.
861
862

863 PROOF. Since patterns are β -normal, $O \geq_{\mathcal{P}} q \cdot F_L$ and from Lemma 5.3 (extended to the set \underline{O}
864 of parallel positions below the fringe), $\underline{u} \xrightarrow[\mathcal{R}ll]{q} s'$ for some s' such that $s = s'\overline{u}^O$. By definition of
865 orthogonal rewriting, $\overline{u}^O \xrightarrow[\beta]{} \gamma$ for some γ such that $t = \underline{u}_O \gamma$.
866

867 We conclude that both $s = s'\overline{u}^O \xrightarrow[\beta]{} s'\gamma$ and, by stability, $t = \underline{u}_O \gamma \xrightarrow[\mathcal{R}ll]{q} s'\gamma$. \square
868
869

870 Definition 5.13. A non-empty set of position $O \subseteq \mathcal{P}os(u)$ is said to be *rigid* in u if there exists
871 $q \leq_{\mathcal{P}} O$ such that $\mathcal{V}ar(u|_O) \subseteq \mathcal{V}ar(u|_q)$.

872 If Q is rigid in u , we can always choose $q = glb(O)$, the greatest lower bound of O w.r.t. $\leq_{\mathcal{P}}$.

873 Note also that a position $o \in \mathcal{P}os(u)$ is a singleton set of rigid positions in u .
874

875 LEMMA 5.14 (LAPOA). Let $s \xleftarrow[\beta]{P} u \xrightarrow[\mathcal{R}ll]{q} t$, where $P \not\leq_{\mathcal{P}} q$. Then $s \xrightarrow[\mathcal{R}ll]{Q'} v \xleftarrow[\beta]{P} t$ where $Q' \geq_{\mathcal{P}} \underline{(P \cup \{q\})}$.
876
877

878 This implies that $s = u'\overline{u}^q$, $v = u'\sigma$ and $t = \underline{u}_q \sigma$ such that $\underline{u}_q \xrightarrow[\beta]{P} u'$, $\overline{u}^q \xrightarrow[\mathcal{R}ll]{} \sigma$.
879

880 Note that $\underline{(P \cup \{q\})}$ are positions in s , hence the condition on Q' makes sense. We could of course
881 conclude $P \not\leq_{\mathcal{P}} Q'$, since it is implied by $Q' \geq_{\mathcal{P}} \underline{(P \cup \{q\})}$, which is therefore more precise.
882

PROOF. We prove a more general result for which $u \xrightarrow[\mathcal{R}ll]{Q} t$, with $P \not\leq_{\mathcal{P}} Q$ and Q is a set of rigid positions of u . We then conclude that $s \xrightarrow[\mathcal{R}ll]{Q'} v$ for some set Q' of positions of s such that $Q' \geq_{\mathcal{P}} (P \cup Q)$.

We prove the result by induction on the set of positions P using the well-founded multiset extension $>_{mul}$ of the size ordering on positions (a set is of course a multiset).

If P or Q is empty, the result is trivial. Otherwise, there are two cases depending whether $\Lambda \in P$.

If $\Lambda \notin P$, then $u = f(\bar{u})$ and $f(\bar{s}) \xleftarrow[\beta]{P} f(\bar{u}) \xrightarrow[\mathcal{R}ll]{Q} f(\bar{t})$ with $s_i \xleftarrow[\beta]{P_i} u_i \xrightarrow[\mathcal{R}ll]{Q_i} t_i$. Since Q is rigid in u , then Q_i is obviously rigid in u_i . Note further that in case two different subsets Q_i and Q_j are non-empty, $glb(Q) = \Lambda$ and $\mathcal{V}ar(u_i|_{Q_i}) \subseteq \mathcal{V}ar(u)$, the latter property being preserved by rewriting. Since $P_i <_{mul} P$, by induction hypothesis, $s_i \xrightarrow[\mathcal{R}ll]{Q'_i} v_i \xleftarrow[\beta]{P_i} t_i$, where $Q'_i \geq_{\mathcal{P}} P_i \cup Q_i$. The orthogonal β -steps can be grouped together into $v \xleftarrow[\beta]{P} t$. For the $\mathcal{R}ll$ -steps, let $Q' = \bigcup_i i \cdot Q'_i$, hence $s = f(\bar{s}) \xrightarrow[\mathcal{R}ll]{Q'} f(\bar{v}) = v$. From $Q'_i \geq_{\mathcal{P}} P_i \cup Q_i$ we deduce $Q' \geq_{\mathcal{P}} P \cup Q$, which concludes this case.

If $\Lambda \in P$, then $u = (\lambda x.M N)$, $P = \{\Lambda\} \uplus P'$, $P' = 1^2 \cdot P_1 \uplus 2 \cdot P_2$ and $Q = 1^2 \cdot Q_1 \uplus 2 \cdot Q_2$, where $P_1, Q_1 \in \mathcal{P}os(M)$ and $P_2, Q_2 \in \mathcal{P}os(N)$ such that $M_w \xleftarrow[\beta]{P_1} M \xrightarrow[\mathcal{R}ll]{Q_1} M_t$ and $N_w \xleftarrow[\beta]{P_2} N \xrightarrow[\mathcal{R}ll]{Q_2} N_t$. Since $P \cup Q = \{\Lambda\}$, we only need to show that $s \xrightarrow[\mathcal{R}ll]{Q'} v \xleftarrow[\beta]{P} t$ for some Q' and v yet to be defined.

There are two cases, depending whether $Q_2 = \emptyset$, the first one being itself split into two:

(1) $Q_2 \neq \emptyset$, a case depicted at Figure 5. Let $w = (\lambda x.M_w N_w) \xleftarrow[\beta]{P'} u \xrightarrow[\mathcal{R}ll]{Q} t$. Since $P' \not\leq_{\mathcal{P}} Q$ and

$P' <_{mul} P$, by induction hypothesis, $w \xrightarrow[\mathcal{R}ll]{Q'} (\lambda x.M_v N_v) \xleftarrow[\beta]{P'} t$. Hence $M_w \xrightarrow[\mathcal{R}ll]{Q'_1} M_v$ and $N_w \xrightarrow[\mathcal{R}ll]{Q'_2} N_v$.

(a) $Q_1 \neq \emptyset$, hence $glb(Q) = \Lambda$. Since Q is a set of rigid positions in u , no variable bound above Q_1 in u can occur in $M|_{Q_1}$. It follows that $x \notin \mathcal{V}ar(M|_{Q_1})$ and furthermore that β -reductions at P_1 do not instantiate terms at Q_1 , and therefore $x \notin \mathcal{V}ar(M_w|_{Q'_1})$. By repeated applications of Lemma 3.11, it follows that $s = M_w \{x \mapsto N_w\} \xrightarrow[\mathcal{R}ll]{Q'} M_v \{x \mapsto N_v\}$, where $Q' := Q'_1 \uplus \{o \cdot Q'_2 : M_w|_o = x\}$ is a set of parallel positions.

(b) $Q_1 = \emptyset$. Then $M_v = M_w$ and $s \xrightarrow[\mathcal{R}ll]{Q'} M_v \{x \mapsto N_v\}$, where $Q' = \{o \cdot Q'_2 : M_w|_o = x\}$ is a set of parallel positions.

Since $t \xrightarrow[\beta]{P' \uplus \{\Lambda\}} v = M_v \{x \mapsto N_v\}$, we are done in both cases.

(2) $Q_2 = \emptyset$, hence $N_t = N$, a case depicted at Figure 6. This time, the variable x may occur below the $\mathcal{R}ll$ -redexes in M , but there are no redexes in N . We split the orthogonal step into three parts:

$$u = (\lambda x.M N) \xrightarrow[\beta]{2 \cdot P_2} (\lambda x.M N_w) = w \xrightarrow[\beta]{\Lambda} M \{x \mapsto N_w\} = \xrightarrow[\beta]{P_1} M_s \{x \mapsto N_w\} = s$$

rewrites at $1^2 \cdot Q_1 \# 2 \cdot P_2$, $(\lambda x.M N_w) \xrightarrow[\mathcal{R}ll]{1^2 \cdot Q_1} (\lambda x.M_t N_w) \xleftarrow[\beta]{2 \cdot P_2} (\lambda x.M_t N) = t$. By stability

Lemma 3.12 used at all positions in Q_1 , $M \{x \mapsto N_w\} \xrightarrow[\mathcal{R}ll]{Q_1} M_t \{x \mapsto N_w\} \xleftarrow[\beta]{\Lambda} (\lambda x.M_t N_w)$. Since $P \not\leq_{\mathcal{P}} Q$, then $P_1 \not\leq_{\mathcal{P}} Q_1$. Besides, since Q is rigid in u and $Q_2 = \emptyset$, then Q_1 must be rigid in M and since substitutions don't capture variables, Q_1 is rigid in $M \{x \mapsto N_w\}$. We can therefore

932 apply the induction hypothesis to $M_s\{x \mapsto N_w\} \xleftrightarrow[\beta]{P_1} M\{x \mapsto N_w\} \xleftrightarrow[\mathcal{R}ll]{Q_1} M_t\{x \mapsto N_w\}$, which
 933
 934 gives $s = M_s\{x \mapsto N_w\} \xleftrightarrow[\mathcal{R}ll]{Q'} M'\{x \mapsto N_w\} \xleftrightarrow[\beta]{P_1} M_t\{x \mapsto N_w\} = v$, where $P \cup Q \geq_P Q'$. Now
 935
 936 $t = (\lambda x.N_t N) \xleftrightarrow[\beta]{2 \cdot P_2 \uplus \Lambda \uplus 1^2 \cdot P_1 = P} v''$ and we are done. □
 937
 938

939 6 CONFLUENCE IN $\lambda\mathcal{F}$

940 We can now address the problem of confluence of a higher-order rewrite theory $\lambda\mathcal{F}$.

941 We assume given a set $\mathcal{R}ll$ of left-linear rewrite rules, and will consider the case where the
 942 relation generated by $\mathcal{R}ll$ is terminating. The other rewrite relations to be considered are $=_\alpha$ and β .
 943 For β , we shall need orthogonal reductions, as previously defined, introduced by Tait under the
 944 name of parallel reductions for showing confluence of the λ -calculus.
 945

946 **THEOREM 6.1.** *Let $(\mathcal{R}ll)$ be a terminating, left-linear rewrite system whose all critical pairs are*
 947 *joinable with $\mathcal{R}ll$. Then $\lambda\mathcal{F}$ is confluent.*

948 **PROOF.** Let $\longrightarrow = \xrightarrow[\mathcal{R}ll]{} \cup \xrightarrow[\beta]{} \circlearrowleft$. The relations \longrightarrow and $\xrightarrow[\lambda\mathcal{F}]{} \circlearrowleft$ verify the assumptions of the following
 949 well-known property: assume $\xrightarrow[\beta]{} \circlearrowleft \subseteq \xrightarrow[\beta]{} \circlearrowleft$ and $\xrightarrow[\beta]{} \circlearrowleft \subseteq \xrightarrow[\beta]{} \circlearrowleft$, then Church-Rosser of $\xrightarrow[\beta]{} \circlearrowleft$ implies
 950 Church-Rosser of $\xrightarrow[\beta]{} \circlearrowleft$. We will therefore apply van Oostrom's decreasing diagram method to the
 951 relation \longrightarrow and conclude that $\xrightarrow[\lambda\mathcal{F}]{} \circlearrowleft$ is confluent. To this end, we use for labels ordered pairs defined as
 952 follows: $\langle 0, u \rangle$ for $u \xrightarrow[\mathcal{R}ll]{} v$; and $\langle 1, \perp \rangle$ for $u \xrightarrow[\beta]{} v$, \perp being a don't care constant. Labels are compared
 953 lexicographically, the first argument in the order on natural numbers, the second in the order $\xrightarrow[\mathcal{R}ll]{} \circlearrowleft$, \perp
 954 being chosen minimal.
 955 We now show that all local peaks have decreasing diagrams.

956 We now show that all local peaks have decreasing diagrams.

957 Let $s \xleftarrow[P]{} u \xrightarrow[Q]{} t$ be an arbitrary local peak, where P, Q are either a set of orthogonal positions (for
 958 β) or a single position (for $\mathcal{R}ll$).

- 959 (1) First, rewrite steps of monotonic relations always commute when $P\#Q$, yielding a DD.

960 We are now left with all peaks for which $\neg(P\#Q)$, which we consider in turn:

- 961 (2) $s \xleftarrow[P]{} u \xrightarrow[Q]{} t$. Orthogonal β -reductions are known to be joinable in at most one step from
 962 each side, hence $s \xrightarrow[\beta]{} v \xrightarrow[\beta]{} t$ for some P', Q', v , a DD.

- 963 (3) $s \xleftarrow[P]{} u \xrightarrow[Q]{} t$ with $i : L \rightarrow R \in \mathcal{R}ll, j \in \mathcal{R}ll$ and $q \geq_P p \cdot F_L$. Lemma 5.4 (LAPRA) yields a
 964 decreasing diagram without facing steps because labels decrease along reductions.

- 965 (4) $s \xleftarrow[P]{} u \xrightarrow[Q]{} t$ with $i : L \rightarrow R \in \mathcal{R}ll, j : G \rightarrow D \in \mathcal{R}ll, q \in p \cdot o$ and $o \in \mathcal{F}Pos(F_L)$. By Lemma 5.6,
 966 there is a critical peak obtained by overlapping G onto L at position o . By assumption, this peak
 967 is joinable with rules of $\mathcal{R}ll$, hence the pair s, t is joinable by the monotonicity lemma 3.11 and
 968 the stability Lemma 3.12. Note that there are no facing steps here, since labels decrease strictly
 969 along $\mathcal{R}ll$ -reductions.

- 970 (5) $s \xleftarrow[O]{} u \xrightarrow[Q]{} t$, where $i : L \rightarrow R \in \mathcal{R}ll$. The proof of this case is shown at Figure 4.
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980

981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029

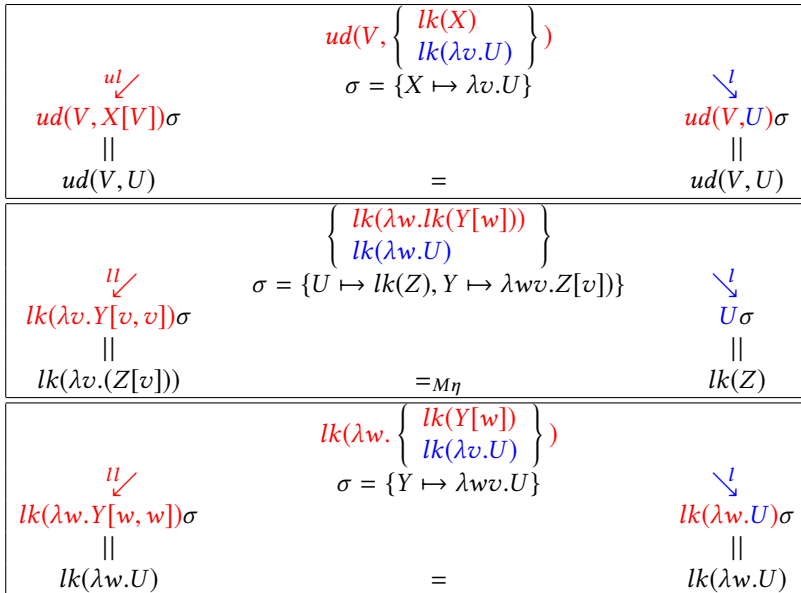
From Lemma 5.10 we have $s \xleftarrow{\otimes R} v \xleftarrow{\otimes Q} u' \xleftarrow{\otimes P} u$ with $O = P \uplus Q \uplus R$ such that $P \# q$, $Q \geq_{\mathcal{P}} q$ and $R <_{\mathcal{P}} q$. Besides, since $u(q) \in \mathcal{F}$, then $q \notin O$ and $Q >_{\mathcal{P}} q$. By commutation we easily get $u' \xrightarrow{q}_i t' \xleftarrow{\otimes P} t$. By Lemma 5.12 we have $v \xrightarrow{q}_i r \xleftarrow{\otimes \geq_{\mathcal{P}} q} t'$ and by Lemma 5.14 we have $s \xrightarrow{i} w \xleftarrow{\otimes R} r$. From Lemma 5.10 again, all three β -steps can be merged into a single orthogonal facing step $w \xleftarrow{\otimes \beta} t$. The step $\xrightarrow{i} w$ can then be linearized, hence we get a DD.

By Lemma 5.1, all cases have been considered, we are therefore done.

Note that the last case in the proof is actually a generalization of (LAPOa) and (LAPOb) to an arbitrary local peak between β - and \mathcal{R} ll-rewrites, which we could have singled out.

Example 6.2. SOL shows the confluence of the theory of global states for the case of simple types with prenex polymorphism. We show below that it is confluent for any type discipline. To this end, we need to show first that it's untyped version is terminating, and then, that the critical pairs are joinable. In the absence of $\beta^{\neq 0}$, first-order termination techniques can do. We are left with verifying the joinability of critical pairs, these computations are presented inside individual boxes. In the upper middle of each box appear two left-hand sides of rules whose superposition is inside braces. The upper left-hand side is displayed in red, the lower one in blue. Next comes the unifier, then the colored right-hand sides, then the reduced right-hand sides, and finally the joinability verification itself, sometimes just an equality test. Colored rule names label the arrows.

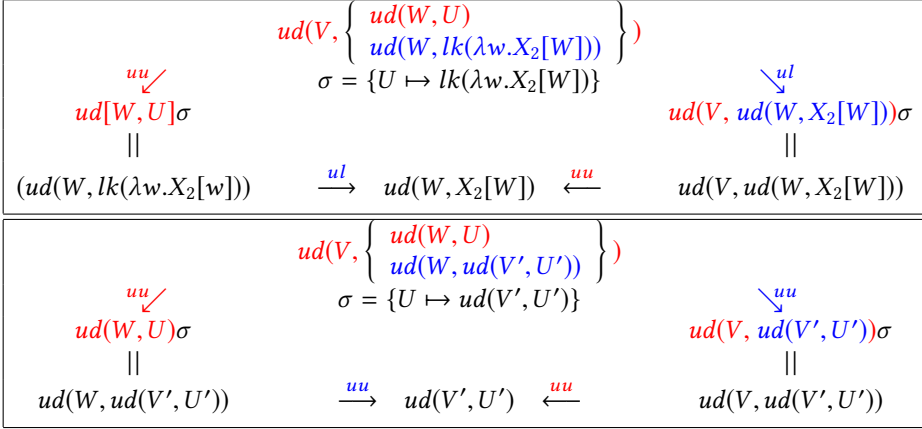
Since the most general unifiers are identical for both choices of rules, we choose the second set. The computations are not identical for both sets, since they will actually depend upon the number of arguments of the meta-variables in the right-hand sides of the rules. Actually, all right-hand sides are the same except for (lu), this will impact four critical pairs exactly.



1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078

$\begin{array}{c} \text{\color{red}ll} \swarrow \\ \text{\color{red}lk}(\lambda v. Y[v, v])\sigma \\ \parallel \\ \text{\color{red}lk}(\lambda v. \text{\color{red}lk}(Y'[v])) \end{array}$	$\text{\color{red}lk}(\lambda w. \left\{ \begin{array}{l} \text{\color{red}lk}(Y[w]) \\ \text{\color{red}lk}(\lambda v. \text{\color{red}lk}(Y'[v])) \end{array} \right\})$ $\sigma = \{Y \mapsto \lambda w v. \text{\color{red}lk}(Y'[v])\}$ $\xrightarrow{\text{\color{red}ll}} \text{\color{red}lk}(\lambda v. Y'[v, v]) \xleftarrow{l}$	$\begin{array}{c} \searrow \text{\color{red}ll} \\ \text{\color{red}lk}(\lambda w. \text{\color{red}lk}(\lambda v. Y'[v, v]))\sigma \\ \parallel \\ \text{\color{red}lk}(\lambda w. \text{\color{red}lk}(\lambda v. Y'[v, v])) \end{array}$
$\begin{array}{c} \text{\color{red}ll} \swarrow \\ \text{\color{red}lk}(\lambda v. Y[v, v])\sigma \\ \parallel \\ \text{\color{red}lk}(\lambda v. \text{\color{red}ud}(v, X[v])) \end{array}$	$\text{\color{red}lk}(\lambda w. \left\{ \begin{array}{l} \text{\color{red}lk}(Y[w]) \\ \text{\color{red}lk}(\lambda v. \text{\color{red}ud}(v, X[v])) \end{array} \right\})$ $\sigma = \{Y \mapsto \lambda w v. \text{\color{red}ud}(v, X[v])\}$ $\xrightarrow{\text{\color{red}lu}} \text{\color{red}lk}(\lambda v. X[v]) =_{M\eta} \text{\color{red}lk}(X) \xleftarrow{l}$	$\begin{array}{c} \searrow \text{\color{red}lu} \\ \text{\color{red}lk}(\lambda w. \text{\color{red}lk}(X))\sigma \\ \parallel \\ \text{\color{red}lk}(\lambda w. \text{\color{red}lk}(X)) \end{array}$
$\begin{array}{c} \text{\color{red}lu} \swarrow \\ \text{\color{red}lk}(X)\sigma \\ \parallel \\ \text{\color{red}lk}(\lambda w. \text{\color{red}lk}(X')) \end{array}$	$\text{\color{red}lk}(\lambda w. \left\{ \begin{array}{l} \text{\color{red}ud}(w, X[w]) \\ \text{\color{red}ud}(V, \text{\color{red}lk}(X')) \end{array} \right\})$ $\sigma = \{V \mapsto w, X \mapsto \lambda w. \text{\color{red}lk}(\lambda v. X'[v])\}$ $\xrightarrow{l} \text{\color{red}lk}(\lambda v. X'[v]) =_{\alpha} \text{\color{red}lk}(\lambda w. X'[w]) \xleftarrow{\text{\color{red}lu}}$	$\begin{array}{c} \searrow \text{\color{red}ul} \\ \text{\color{red}lk}(\lambda w. \text{\color{red}ud}(V, X'[V]))\sigma \\ \parallel \\ \text{\color{red}lk}(\lambda w. \text{\color{red}ud}(w, X'[w])) \end{array}$
$\begin{array}{c} \text{\color{red}lu} \swarrow \\ \text{\color{red}lk}(\lambda w. X[w])\sigma \\ \parallel \\ \text{\color{red}lk}(\lambda w. \text{\color{red}lk}(\lambda w. \text{\color{red}ud}(V, W))) \end{array}$	$\text{\color{red}lk}(\lambda w. \left\{ \begin{array}{l} \text{\color{red}ud}(w, X[w]) \\ \text{\color{red}ud}(U, \text{\color{red}ud}(V, W)) \end{array} \right\})$ $\sigma = \{U \mapsto w, X \mapsto \lambda w. \text{\color{red}ud}(V, W)\}$ $\xrightarrow{\text{\color{red}ll}}$	$\begin{array}{c} \searrow \text{\color{red}uu} \\ \text{\color{red}lk}(\lambda w. \text{\color{red}ud}(V, W))\sigma \\ \parallel \\ \text{\color{red}lk}(\lambda w. \text{\color{red}ud}(V, W)) \end{array}$
$\begin{array}{c} \text{\color{red}ul} \swarrow \\ \text{\color{red}ud}(V, X_1[V])\sigma \\ \parallel \\ \text{\color{red}ud}(V, \text{\color{red}ud}(V, X_2[V])) \end{array}$	$\text{\color{red}ud}(V, \left\{ \begin{array}{l} \text{\color{red}lk}(\lambda v. X_1[v]) \\ \text{\color{red}lk}(\lambda v. \text{\color{red}ud}(v, X_2[v])) \end{array} \right\})$ $\sigma = \{X_1 \mapsto \lambda v. \text{\color{red}ud}(v, X_2[v])\}$ $\xrightarrow{\text{\color{red}uu}} \text{\color{red}ud}(V, X_2[V]) \xleftarrow{\text{\color{red}ul}}$	$\begin{array}{c} \searrow \text{\color{red}lu} \\ \text{\color{red}ud}(V, \text{\color{red}lk}(\lambda v. X_2[v]))\sigma \\ \parallel \\ \text{\color{red}ud}(V, \text{\color{red}lk}(\lambda v. X_2[v])) \end{array}$
$\begin{array}{c} \text{\color{red}ul} \swarrow \\ \text{\color{red}ud}(V, X_1[V])\sigma \\ \parallel \\ \text{\color{red}ud}(V, \text{\color{red}lk}(\lambda w. X_2[w, V])) \end{array}$	$\text{\color{red}ud}(V, \left\{ \begin{array}{l} \text{\color{red}lk}(\lambda v. X_1[v]) \\ \text{\color{red}lk}(\lambda v. \text{\color{red}lk}(\lambda w. X_2[w, v])) \end{array} \right\})$ $\sigma = \{X_1 \mapsto \lambda v. \text{\color{red}lk}(\lambda w. X_2[w, v])\}$ $\xrightarrow{\text{\color{red}ul}} \text{\color{red}ud}(V, X_2[V, V]) \xleftarrow{\text{\color{red}ul}}$	$\begin{array}{c} \searrow \text{\color{red}ll} \\ \text{\color{red}ud}(V, \text{\color{red}lk}(\lambda w. X_2[w, w]))\sigma \\ \parallel \\ \text{\color{red}ud}(V, \text{\color{red}lk}(\lambda w. X_2[w, w])) \end{array}$

1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127



Hence, all critical pairs are joinable, or joinable modulo $M\eta$ for two of them. It follows that the theory of global states for a single location preserves confluence of the β -rule in the pure λ -calculus.

Note finally that most of these critical pairs are not development closed, since they need be joined from both sides.

6.1 Relationship to Nipkow’s higher-order rewriting

Nipkow’s rewriting assumes terms to be simply typed, but it can be easily extended to other typing disciplines. The major requirement is indeed that β -reduction is strongly normalizing as well as η -expansion. The latter is obtained by restricting its application to functionally typed terms (which can be obtained in our case by controlling the arity of expressions).

Assuming a subset \mathcal{T} of the set of terms that satisfies these assumptions, we denote by $u \downarrow_\beta$, $u \uparrow^\eta$ and $u \uparrow_\beta^\eta$ the β -normal form, the η -expanded form and the β -normal η -expanded form, respectively, possibly omitting indices and exponents when convenient.

A rule “à la Nipkow” assumes η -expanded left-hand side patterns and η -expanded right-hand sides as well as fully applied meta-variables of arity zero. To have both Nipkow’s rewriting relation and ours defined in our setting, the pre-redex $(X \bar{x})$ in a Nipkow’s pattern will correspond in our syntax to the pre-redex $X[\bar{x}]$ in which $ar(X) = |\bar{x}|$. It follows that a rule will have two different writings dubbed Klop and Nipkow, respectively. We will denote by \mathcal{R}_{kp} the set of higher-order Klop rules, corresponding to a set \mathcal{R}_{nw} of Nipkow rules, which must therefore be in η -expanded form.

In the Nipkow case, because meta-variables have arity zero, Klop’s notion of substitution is nothing but the usual higher-order substitution. The meaning of the same expression $L\sigma$ for some left-hand side of rule L will therefore depend whether the rule $L \rightarrow R$ belongs to \mathcal{R}_{nw} or \mathcal{R}_{kp} : the equality $u = L\sigma$ when L is a Klop left-hand side of rule becomes $u =_{\beta^0} L'\sigma'$ for the corresponding Nipkow left-hand side of rule L' . The same applies to unification of left-hand sides.

We now make these remarks formal:

Definition 6.3. $u \xrightarrow[p]{L \rightarrow R} v$ for $L \rightarrow R \in \mathcal{R}_{nw}$ iff $u = u \downarrow_\beta$, $u \downarrow_p =_{\beta^0} L\sigma$ for some β -normal η -expanded substitution σ and $v = u[R\sigma]_p \downarrow$.

We write $u \xrightarrow[p]{\mathcal{R}_{nw}} v$ when $u \xrightarrow[p]{L \rightarrow R} v$ for some $L \rightarrow R \in \mathcal{R}_{nw}$. As it is known that η -expanded forms are closed under β -reduction and substitution, we have:

LEMMA 6.4. If $u \xrightarrow[\mathcal{R}_{nw}]{p} v$ then $u = u\Downarrow$ and $u \xrightarrow[\mathcal{R}_{kp}]{p} \xrightarrow[\beta]{} v = v\Downarrow$.

LEMMA 6.5. For all $u \in \mathcal{T}$ such that $u \xrightarrow[\mathcal{R}_{kp}]{p} v$ and $v \in \mathcal{T}$ then $u\Downarrow \xrightarrow[\mathcal{R}_{nw}]{} v\Downarrow$.

PROOF. We have $u = u[L\sigma]_p = \bar{u}^p \{Z \mapsto \lambda \bar{z}.L\sigma\} = \bar{u}^p \gamma$ and $v = u[R\sigma]_p = \bar{u}^p \{Z \mapsto \lambda \bar{z}.R\sigma\} = \bar{u}^p \theta$ with $\gamma \xrightarrow[\mathcal{R}]{} \theta$. Since L is a pattern $(L\sigma)\Downarrow = L(\sigma\Downarrow)$ and since it cannot be an abstraction we have,
 $u\Downarrow = ((\bar{u}^p \gamma)\Downarrow)\Uparrow = (\bar{u}^p \Downarrow (\gamma\Downarrow))\Uparrow \xrightarrow[\mathcal{R}_{nw}]{} (\bar{u}^p \Downarrow \{Z \mapsto \lambda \bar{z}.R(\sigma\Downarrow)\})\Downarrow = ((\bar{u}^p \{Z \mapsto \lambda \bar{z}.R\sigma\})\Downarrow)\Uparrow = v\Downarrow$ \square

COROLLARY 6.6. If $u \xrightarrow[\beta \eta \mathcal{R}_{kp}]{} v$ then $u\Downarrow \xrightarrow[\mathcal{R}_{nw}]{} v\Downarrow$.

LEMMA 6.7. Assume a rewrite system \mathcal{R} such that $\xrightarrow[\beta]{} \cup \xrightarrow[\mathcal{R}_{kp}]{} is Church-Rosser. Then \xrightarrow[\mathcal{R}_{nw}]{} is Church-Rosser on \beta-normal \eta-expanded terms of \mathcal{T}.$

PROOF. Assume $u \xleftarrow[\mathcal{R}_{nw}]{} v$. Then by Lemma 6.5, $u = u\Downarrow$, $v = v\Downarrow$ and $u \xleftarrow[\beta \eta \mathcal{R}_{kp}]{} v$. By assumption, $u (\xrightarrow[\mathcal{R}_{kp}]{} \cup \xrightarrow[\beta]{})^* (\xleftarrow[\mathcal{R}_{kp}]{} \cup \xleftarrow[\beta]{})^* v$ and by Corollary 6.6, $u \xrightarrow[\mathcal{R}_{nw}]{} \xleftarrow[\mathcal{R}_{nw}]{} v$. \square

THEOREM 6.8. Assume \mathcal{R} is a left-linear system such that all critical pairs are joinable using $\xrightarrow[\mathcal{R}_{kp}]{} and \mathcal{R} is strongly normalizing. Then $\xrightarrow[\mathcal{R}_{nw}]{} is Church-Rosser on \beta-normal \eta-expanded terms of \mathcal{T}.$$

PROOF. $\xrightarrow[\mathcal{R}]{} satisfies the hypotheses of Theorem 6.1. We conclude that \xrightarrow[\beta]{} \cup \xrightarrow[\mathcal{R}]{} is Church-Rosser and by Lemma 6.7, \xrightarrow[\mathcal{R}_{nw}]{} is too on \beta-normal \eta-expanded terms of \mathcal{T}.$ \square

This result is of course not really surprising, since it is known to hold for various type systems for which β -reductions terminate, in which case \mathcal{T} is the set of typable terms. What's new here is only that we abstract from a particular typing discipline.

7 CONCLUSION

Confluence of first-order rewrite rules is well understood, in both the terminating and non-terminating cases. Confluence of left-linear higher-order rules on simply-typed λ -terms is well understood too [12]. This is true as well of confluence of first-order rules in presence of β -reductions for any type discipline for which β is terminating [4].

In this paper, we have described a condition on critical pairs which ensures preservation of confluence in the untyped λ -calculus by a set $\mathcal{R}ll$ of rewrite rules whose left-hand sides are left-linear patterns: if $\mathcal{R}ll$ is terminating and the critical pairs of $\mathcal{R}ll$ are joinable by rules of $\mathcal{R}ll$. The β -rule itself, on the other hand, cannot be used to join the critical pairs. In that case, confluence can be obtained by joining *nested critical pairs*, as will be shown in the second paper of this series. This other result does not subsume the present one, since nested critical pairs may be infinitely many, as is the case with the theory of global states for a unique location [6], which is shown here to preserve confluence of β -reductions by computing its finitely many critical pairs.

In order to define critical pairs, we had to unify left-linear patterns, where patterns are specific untyped λ -terms whose definition ensures that erasing types from a simply-typed pattern in Miller's sense yields a pattern in our sense. Untyped patterns enjoy most general unifiers, in the same way as Miller's patterns do. Note that unification here looks purely syntactic, thanks to a definition of substitution which incorporates β -reductions. The linearity restriction should not be essential: having multiple occurrences of meta-variables requires using a merge rule in the unification case,

1177 and checking terms for equality in the matching case. This will be carried out in details in the third
 1178 paper of this series, in which we shall also discuss the similarities and differences of our algorithm
 1179 with those of Miller [13] and Nipkow [12].

1180 One may wonder why we did not consider a well-known setting, like Klop's combinatory
 1181 reduction systems [11] or Nipkow's higher-order rewriting, or van Oostrom's higher-order rewriting
 1182 systems [20], and then encode our notion of higher-order rewriting within their's. One main reason
 1183 is that we always insist, in DEDUKTI, in using shallow encodings, hence do not want to encode the
 1184 λ -calculus itself as a higher-order calculus in such a setting. Further, our notion of meta-variable
 1185 has a fixed arity but may have missing arguments, which is unusual. Although one could fear that
 1186 the present setting becomes too specific for a wide application, we believe that this is not the
 1187 case, and that it can be used to show confluence of rewrite rules in other dependent type theories
 1188 without difficulty, as well as for other, related rewrite relations, as we have shown with Nipkow's
 1189 higher-order rewriting.

1190 One may also wonder whether considering parallel higher-order critical pairs could improve our
 1191 results. The difficulty here is that one of the decreasing diagrams for free, Lemma 5.4, breaks down.
 1192 It can of course be repaired, to the price of imposing that meta-variables do not occur embedded
 1193 in one another in the right-hand sides of the rules. This restriction looks of course very strong.
 1194 However, any expression such as $X[Y]$ can be transformed into $(X Y)$, hence eliminating this
 1195 embedding. There is of course a general transformation that will eliminate all embeddings, making
 1196 the use of parallel rewriting (and therefore parallel critical pairs) look attractive. The problem
 1197 however, is that right-hand sides such as $(X Y)$ may result in the use of β -steps to join the critical
 1198 pairs, hence the joinability diagrams would not be decreasing. This may or may not happen, of
 1199 course. It is certainly possible to exhibit examples for which this transformation would work. We
 1200 have not encountered such a natural example so far. The next paper in this series will therefore
 1201 address directly an even more general left-linear case, by using orthogonal higher-order rewriting.

1202 The case of non-left-linear rules is not touched at all here, it is indeed much more difficult since
 1203 adding such rules to the untyped λ -calculus results, in general, in loosing confluence, as shown
 1204 by Klop [11]. We however show in another forthcoming paper that all Klop's counter-examples
 1205 preserve confluence of the set of β -strongly-normalizing λ -terms, hence showing a way to get
 1206 around this difficulty. Finally, mixing left-linear rules with right-linear ones, is a problem which is
 1207 important to us, because encodings of complex type theories in DEDUKTI are not purely left-linear,
 1208 and of course not purely right-linear either. We do not know yet whether we can obtain meaningful
 1209 results for this combination.

1210 **Acknowledgments:** to Gilles Dowek for many discussions, Jiaxiang Liu for a chary reading,
 1211 and Vincent van Oostrom for his many suggestions and corrections to an earlier draft.

1213 REFERENCES

- 1214 [1] Ali Assaf, Guillaume Burel, Raphaël Cauderlier, Gilles Dowek, Catherine Dubois, Frédéric Gilbert, Pierre Halma-
 1215 grand, Olivier Hermant, and Ronan Saillard. Dedukti: a Logical Framework based on the lambda-pi-Calculus Modulo
 1216 Theory. draft, INRIA, 2019.
- 1217 [2] Hendrik Pieter Barendregt. *The lambda calculus : its syntax and semantics*. Studies in logic and the foundations of
 1218 mathematics. North-Holland, Amsterdam, New-York, Oxford, 1981.
- 1219 [3] Jesper Cockx, Nicolas Tabareau, and Théo Winterhalter. How to tame your rewrite rules, 2018. Draft.
- 1220 [4] Daniel J. Dougherty. Adding algebraic rewriting to the untyped lambda calculus. *Inf. Comput.*, 101(2):251–267, 1992.
- 1221 [5] Gilles Dowek at all. The Dedukti system, 2016. Available from <http://dedukti.gforge.inria.fr/>.
- 1222 [6] Gilles Dowek, Jean-Pierre Jouannaud and Jiaxiang Liu. Confluence in untyped higher-order theories. draft hal-, INRIA,
 1223 january 2019. Full version of a work presented at HOR 2016.
- 1224 [7] Healfdene Goguen. The metatheory of UTT. In Peter Dybjer, Bengt Nordström, and Jan M. Smith, editors, *Types for*
 1225 *Proofs and Programs, International Workshop TYPES'94, Båstad, Sweden, June 6-10, 1994, Selected Papers*, volume 996 of

- 1226 *Lecture Notes in Computer Science*, pages 60–82. Springer, 1994.
- 1227 [8] Makoto Hamana. How to prove your calculus is decidable: practical applications of second-order algebraic theories
1228 and computation. *PACMPL*, 1(ICFP):22:1–22:28, 2017.
- 1229 [9] J. R. Hindley. An abstract form of the Church-Rosser theorem. i. *J. Symb. Log.*, 34(4):545–560, 1969.
- 1230 [10] Jean-Pierre Jouannaud and Jiaxiang Liu. From diagrammatic confluence to modularity. *Theor. Comput. Sci.*, 464:20–34,
1231 2012.
- 1232 [11] Jan Willem Klop. *Combinatory reduction systems*. PhD thesis, CWI tracts, 1980.
- 1233 [12] Richard Mayr and Tobias Nipkow. Higher-order rewrite systems and their confluence. *Theoretical Computer Science*,
1234 192:3–29, 1998.
- 1235 [13] Dale Miller. A logic programming language with lambda-abstraction, function variables, and simple unification.
1236 *Journal of Logic and Computation*, 1(4):497–536, 1991.
- 1237 [14] Maxwell H. A. Newman. On theories with a combinatorial definition of ‘equivalence’. *Ann. Math.*, 43(2):223–243, 1942.
- 1238 [15] Mitsuhiro Okada. Strong normalizability for the combined system of the typed lambda calculus and an arbitrary
1239 convergent term rewrite system. In Gaston H. Gonnet, editor, *Proceedings of the ACM-SIGSAM 1989 International
1240 Symposium on Symbolic and Algebraic Computation, ISSAC '89, Portland, Oregon, USA, July 17-19, 1989*, pages 357–363.
ACM, 1989.
- 1241 [16] Gordon D. Plotkin and John Power. Algebraic operations and generic effects. *Applied Categorical Structures*, 11(1):69–94,
1242 2003.
- 1243 [17] Vincent van Oostrom. Confluence by decreasing diagrams. *Theor. Comput. Sci.*, 126(2):259–280, 1994.
- 1244 [18] Vincent van Oostrom. Developing developments. *Theor. Comput. Sci.*, 175(1):159–181, 1997.
- 1245 [19] Vincent van Oostrom. Confluence by decreasing diagrams converted. In Voronkov A., editor, *RTA*, volume 5117 of
1246 *Lecture Notes in Computer Science*, pages 306–320. Springer, 2008.
- 1247 [20] Vincent van Oostrom and Femke van Raamsdonk. Comparing combinatory reduction systems and higher-order
1248 rewrite systems. In Jan Heering, Karl Meinke, Bernhard Möller, and Tobias Nipkow, editors, *Higher-Order Algebra,
1249 Logic, and Term Rewriting, First International Workshop, HOA '93, Amsterdam, The Netherlands, September 23-24, 1993*,
1250 *Selected Papers*, volume 816 of *Lecture Notes in Computer Science*, pages 276–304. Springer, 1993.

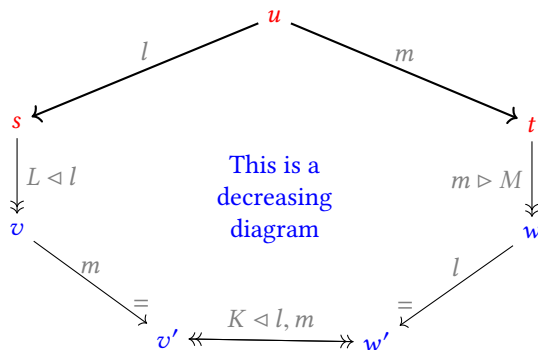


Fig. 2. Decreasing diagram

1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323

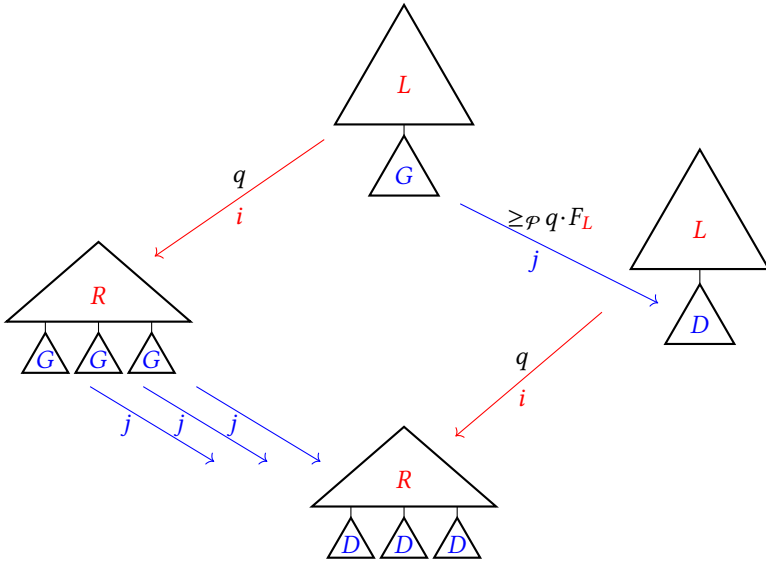


Fig. 3. Ancestor peaks in rewrite theories. L, G stand for terms rewriting to R, D , using a red rule in $\mathcal{R} \cup \beta$ and a blue rule in $\mathcal{R} \cup \beta$.

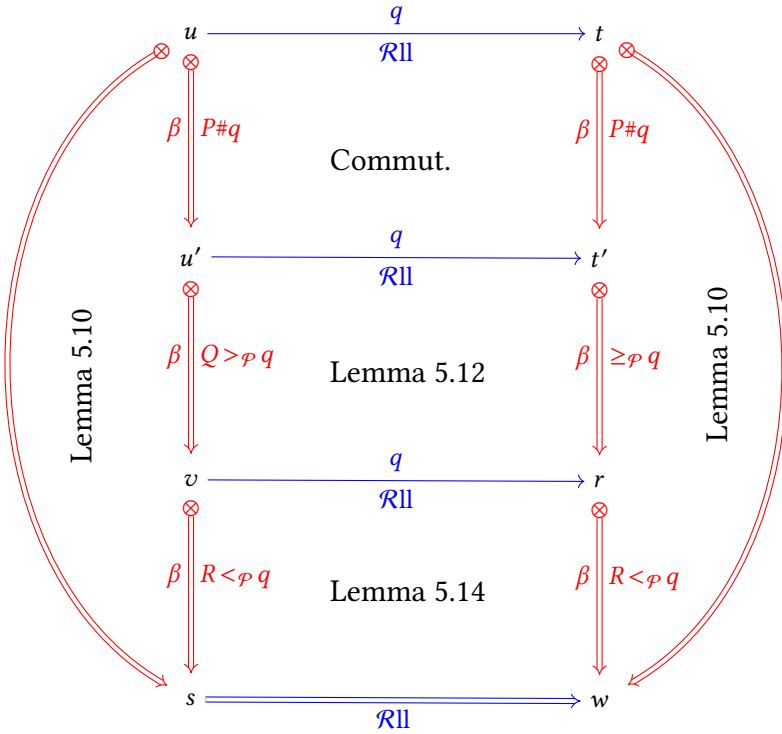


Fig. 4. Construction of a decreasing diagram for heterogeneous local peaks.

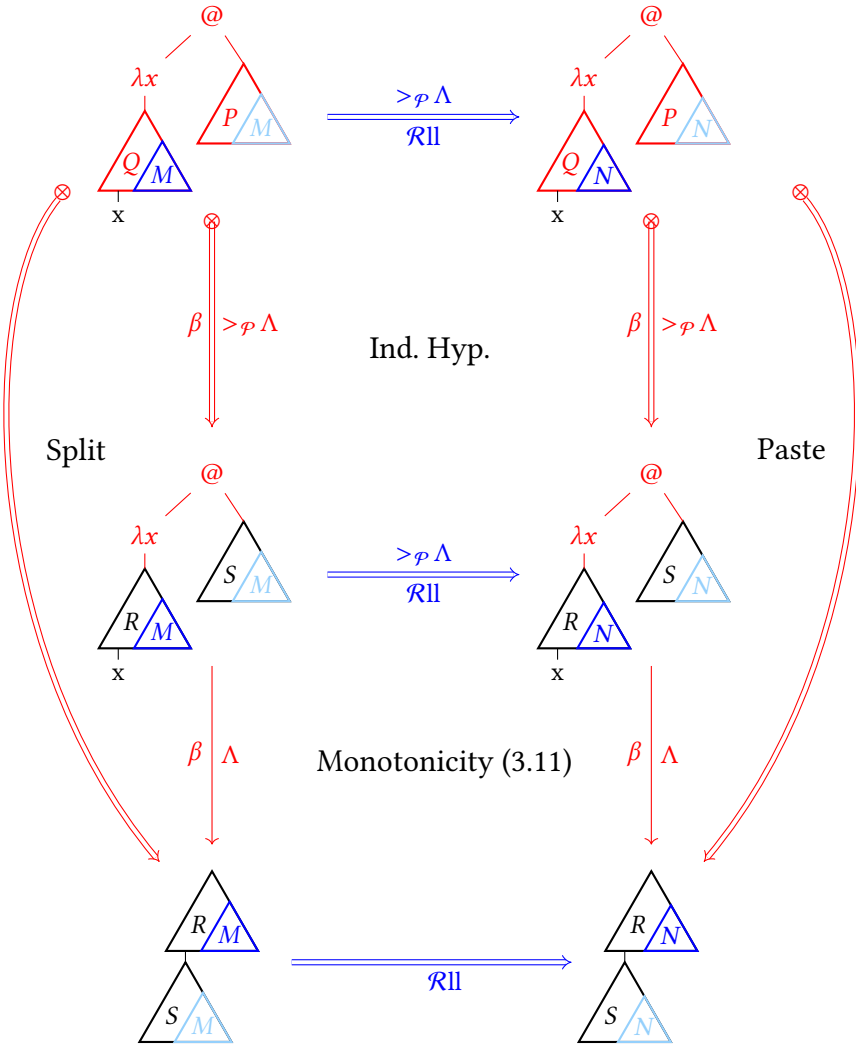


Fig. 5. Construction of a decreasing diagram for peak: Lemma 5.14

1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372

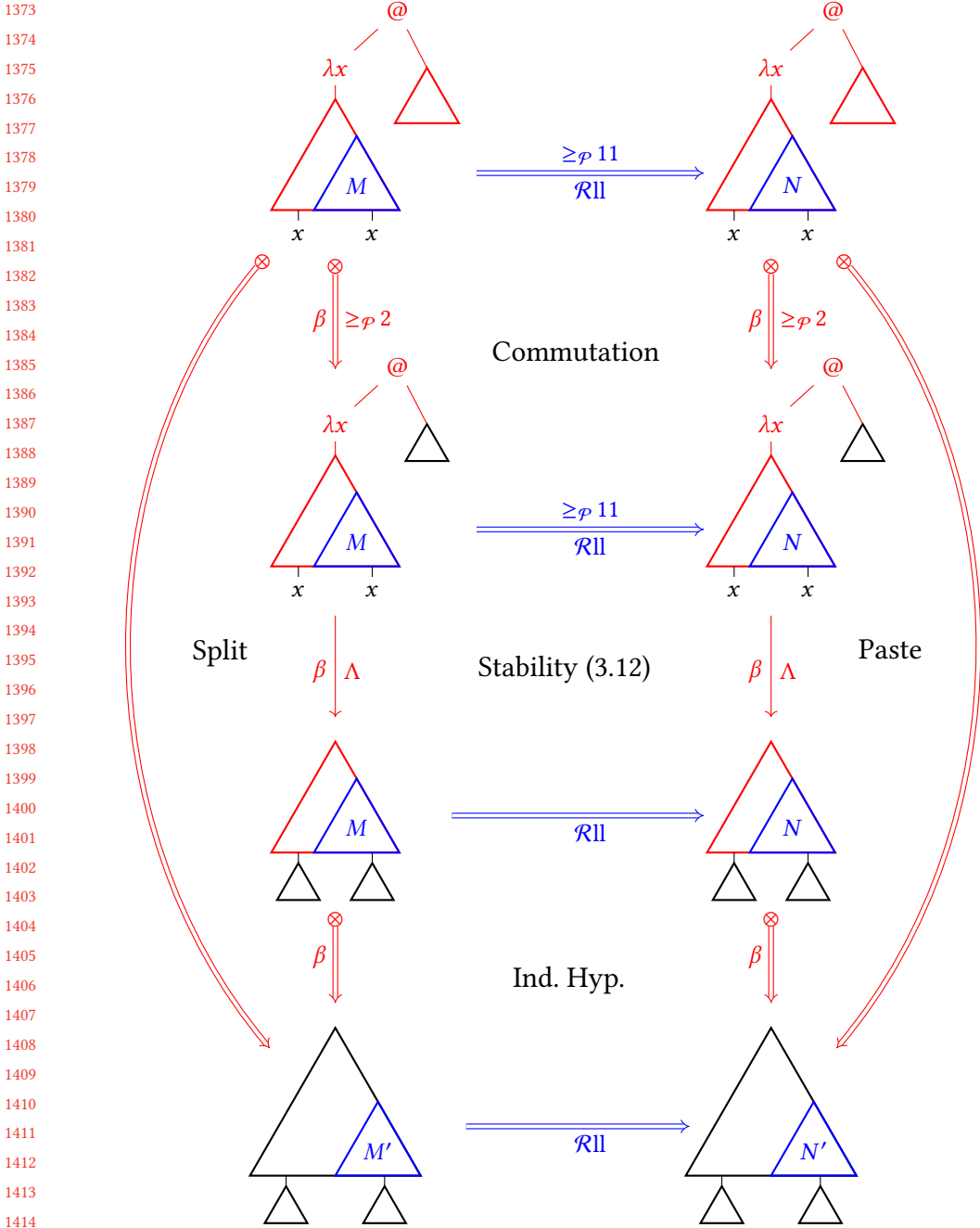


Fig. 6. Construction of a decreasing diagram for peak: Lemma 5.14