5

Redex Families and Optimality

In 1978, Lévy introduced the notion of *redex family* in the λ-calculus with the aim of formally capturing an *intuitive* idea of *optimal sharing* between 'copies' of the same redex. In order to satisfy his notion of family, Lévy proposed several alternative definitions inspired by different perspectives and proved their equivalence.

The most abstract approach to the notion of family (see [Lév78, Lév80]) is the so called *zig-zag* relation. In this case, duplication of redexes is formalized as residuals modulo permutations. In particular, a redex u with history σ (notation σu) is a *copy* of a redex v with history ρ iff $\rho v \leq \sigma u$ (*i.e.*, there exists τ such that $\sigma = \rho\tau$ up to a notion of permutation of redexes, and u is a residual of v after τ). The family relation $\simeq$ is then the symmetric and transitive closure of the copy relation. Now, let us draw reduction arrows downwards. Pictorially, the family reduction gives rise to an alternate sequence of descending and ascending reduction arrows. This is the reason why it is also known as the 'zig-zag' relation.

Another approach is that of considering the *causal history* of redexes. Intuitively, two redexes can be 'shared' if and only if they have been 'created in the same way' (or, better, their *causes* are the same). This is formalized by defining an *extraction relation* over redexes (with history) σu, which throws away all the redexes in σ that have not been relevant for the creation of u. The canonical form we obtain at the end of this process essentially expresses the causal dependencies of u along the derivation (we may deal with causal *chains* instead of *partial orders* since *only standard derivations* are considered).

The most 'operational' approach to the family relation is based on a suitable labeled variant of the λ-calculus [Lév78]. The idea of labels is essentially that of marking the 'points of contact' created by reductions.

In particular, labels grow along the reduction, keeping a trace of its history. Two redexes are in a same family if and only if their labels are identical.

The equivalence between zig-zag and extraction is not particularly problematic (see [Lév80]). On the contrary, the proof of the equivalence between extraction (or zig-zag) and labeling is much more difficult (see for instance the long proof that Lévy gave in [Lév78]). For this reason, we shall postpone its proof to Chapter 6, after the introduction of the so-called *legal paths*. In fact, given a term T, legal paths will give us a complete characterization in terms of paths (virtual redexes) in T of the redexes generated along the reduction of T. Intuitively, a legal path for a redex R will be a suitable composition of subpaths connecting those redexes required for the creation of R. Any legal path φ will correspond to a reduction ρ composed of redexes in φ only; moreover, the redex with history ρR will be in normal form with respect to the extraction relation. The main issue of section 6.2 of Chapter 6, will be to prove that the labels generated along the reduction of T 'are' indeed legal paths in T, and vice versa.

5.1 Zig-Zag

Let $\Delta = \lambda x.(x\,x)$, $F = \lambda x.(x\,y)$, $I = \lambda x.x$. Consider the possible reductions of $M = (\Delta\,(F\,I))$, represented in Figure 5.1.

Intuitively, there are just three kinds of redexes in this example: R, S and T. In the case of R and S, the common nature of these redexes is clear, for all R_i and S_i are residuals of R and S. The case of T is more complex. In fact, T, T_1 and T_2 are not residuals of any redex in M: they are just created by the immediately previous reduction. The two redexes T_1 and T_2 clearly look sharable (and they are indeed 'shared' as the unique redex T in the innermost reduction on the left). However, the only way to establish some formal relation between these redexes is by closing residuals downwards. In fact, T_1 and T have a common residual T_3 inside the term $((I\,y)\,(I\,y))$; similarly, T_2 and T have a common residual T_4 in the same term. By transitive closure, we can thus conclude the 'common nature' of T_1 and T_2: in a sense, T_1 is 'the same' redex as T_3, which is the same as T, which in turn is the same as T_4 which is the same as T_2. So, the idea of 'closing residuals downwards' seems the right extension allowing connection of redexes with a common origin, but without any common ancestor.

The formalization of the previous intuition requires, however, some

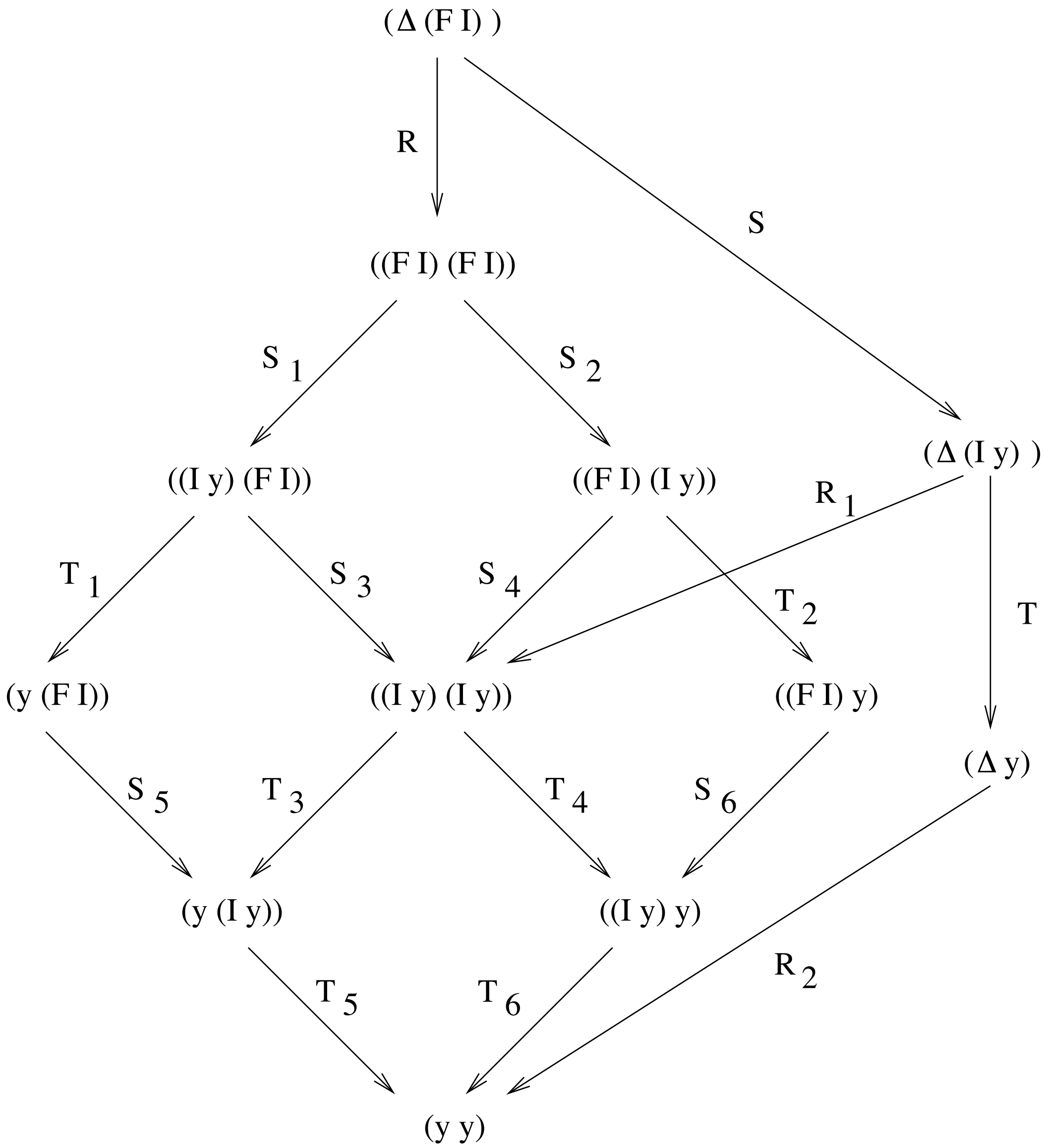


Fig. 5.1. $\Delta = \lambda x.\,(x\,x), F = \lambda x.\,(x\,y), I = \lambda x.\,x.$

care. In particular, let us start by noticing that T_1 and T_2 should not be connected if the initial expression is $((F\,I)\,(F\,I))$, instead of M. The intuitive reason is that, while we want to preserve the sharing 'inherent' in the initial λ-term, we are not interested in recognizing common subexpressions generated along the reduction—the equivalence of such new subexpressions should be considered as a mere syntactical coincidence. For instance, although when reducing any of the two redexes in $(I\,(I\,x))$ we get the 'same' term, the term $(I\,x)$ obtained reducing the outermost redex and the term $(I\,x)$ obtained reducing the innermost redex should be considered distinct. The reasons should be computationally clear: to look for common sharable subexpressions would be too expensive in any practical implementation. Furthermore, even though we could imagine an optimization step based on a preprocessing recognizing common subexpressions in the initial term, to run such an optimizing algorithm at run-time would be infeasible.

A main consequence of the latter assumption is that the relation we are looking for will not really be a relation over redexes, it will be rather relativized with respect to the reduction of some initial expression. In other words, we shall have to consider pairs composed of a redex R together with the reduction ρ that generated it. Namely, given a derivation $T \overset{\rho}{\twoheadrightarrow} T' \overset{R}{\to}$, we shall say that ρ is the *history* of the redex R in T'. As a consequence, any redex with history ρR determines a unique initial term T with respect to which we are considering sharable redexes, and two redexes with history will be comparable only when they start at the same initial term.

Let us come back to the see how the previous considerations apply to our previous examples. Since we relativized redexes with respect to an initial expression and we equipped them with a history, redexes are no longer identified only by their syntactical position in the term. For instance, the term $M = (I\,(I\,x))$ has two redexes R and S. By firing any of them we obtain the term $(I\,x)$, in which we have a unique redex T. Since T is a residual of both R and S we could be tempted to conclude that R and S are connected by our 'sharing' relation. On the contrary, for we do not want to relate distinct redexes in the initial term, the equivalence of the result reducing R or S is merely incidental. In fact, introducing histories, the two reductions respectively give RT and ST. Thus, the two redexes RT and ST can be related via our sharing relation, only if their histories R and S are related via such a relation. That is, only if M is some partial result of the computation of another term N, and we can find two reductions ρ and σ of N such that ρR and σS might be equated applying zig-zag or the extraction relation. This is indeed the case for the term $((I\,y)\,(I\,y))$ generated along the reduction in Figure 5.1. The two redexes T_3 and T_4 do not have any common ancestor. Nevertheless, zig-zag will relate them (see Example 5.1.9) and the extraction relation will associate to any ρT_3 and any σT_4 the same redex with history ST (see Example 5.2.3 and Exercise 5.2.5).

5.1.1 Permutation equivalence

Before to give the definition of families, we present some standard results of λ-calculus. For an unabridged presentation of λ-calculus and for the proof of the previous results, we refer the reader to Hindley and Seldin's book [HS86], even though the most complete source for the syntax properties of λ-calculus is definitely Barendregt's book [Bar84]. Nevertheless, especially for readers unfamiliar with λ-calculus, we suggest starting with Hindley and Seldin. We also remark that the proofs of the following results are actually subsumed by the results on the labeled calculus that we give in section 5.3.

The first notions we need are those of *residual* of a redex (note that in the proof of correctness we have already met this concept using graphs) and *permutation* of a reduction. To introduce them, let us recollect what we did learn by analyzing the example of $(\mathsf{I}\,(\mathsf{I}\,y))$. The relevant point was that the syntactical equivalence of the redexes RT (T with history R) and ST (T with history S) is incidental, for the redex T in $(\mathsf{I}\,y)$ is respectively a residual of S w.r.t. R and a residual of R w.r.t. S. Abstracting from the contingent syntax of λ-calculus that equates the result of the two reductions above, there is clearly no reason to equate RT and ST.

The previous point is settled by introducing a permutation operation on redexes such that two reductions $\mathsf{T} \overset{\rho}{\twoheadrightarrow} \mathsf{T}'$ and $\mathsf{T} \overset{\sigma}{\twoheadrightarrow} \mathsf{T}'$ are permutation equivalent only when ρ is a suitable permutation of the redexes in σ. By the way, in this framework the word permutation must be interpreted in a wide sense, for the contraction of a redex might cause duplication of another redex following it; thus, permuting a reduction, its length might shrink or expand. In the previous example, since R and S are distinct redexes of the initial term, the corresponding reductions cannot be equated by permutation. Furthermore, in such a formal setting we should not use the name T to denote the redex of $(\mathsf{I}\,y)$; taking into account that reducing R the redex T is what remains of the redex S, while reducing S the redex T is what remains of R, we should rather use the names S/R (the residual of S after R) and S/R. It is then intuitive that in order to get two equivalent reductions, the reductions R and S should be completed reducing S/R and R/S, respectively. Namely, while RS/R and SR/S will not be equated as redexes with history, the two reductions RS/R and SR/S are obviously equivalent by permutation.

To denote the residual of a redex, let us assume that we can mark some redexes of a λ-term by underlining them and associating a name to each underlining. For instance, let us consider again the term of Figure 5.1

$$\underline{(\Delta\,\underline{(\mathsf{F}\,\mathsf{I})}_{\mathsf{S}})}_{\mathsf{R}}$$

we have underlined the redexes R and S of $(\Delta\,(\mathsf{F}\,\mathsf{I}))$ using the names R and S (we say that the redexes are underlined by R and S, respectively).

Let us assume that β-reduction preserves marking and naming. It is intuitive that, after a reduction, what is left of a redex marked by a name R, say its residual, is formed of all the redexes of the result with the same marking R.

Example 5.1.1 For instance,

$$\underline{(\Delta\,\underline{(\mathsf{F}\,\mathsf{I})}_{\mathsf{S}})}_{\mathsf{R}} \xrightarrow{\mathsf{R}} (\underline{(\mathsf{F}\,\mathsf{I})}_{\mathsf{S}}\,\underline{(\mathsf{F}\,\mathsf{I})}_{\mathsf{S}}) \xrightarrow{\mathsf{S}_1} ((\mathsf{I}\,y)\,\underline{(\mathsf{F}\,\mathsf{I})}_{\mathsf{S}}) \xrightarrow{\mathsf{S}_3} ((\mathsf{I}\,y)\,(\mathsf{I}\,y))$$

is a marked reduction taken from the example in Figure 5.1.

The previous example also shows that when a marked redex is fired there is no trace of its underlining in the result, which corresponds to the fact that after its contraction a redex has no residual.

Definition 5.1.2 (residual) Let $\rho : \mathsf{M} \twoheadrightarrow \mathsf{N}$. Let us assume that R is the only redex of M marked by an underlining with name R. The set R/ρ of the redexes of N marked by an underlining with name R is the *residual* of R under ρ. The redex R is the *ancestor* of any redex $\mathsf{R}' \in \mathsf{R}/\rho$, while any of these R' is a *residual redex* of R.

The previous definition gives an example of how to use marking. In general, given a reduction $\rho : \mathsf{M} \twoheadrightarrow \mathsf{N}$, we associate to ρ a corresponding marked reduction assuming that each redex of M is underlined using its name. As a consequence, two redexes R and S of M will always be underlined using distinct names. Let us however remark that this is no longer true for the result of a reduction (see the example above).

Let $\mathcal{F}$ be a set of redexes of a λ-term M. A reduction $\rho : \mathsf{M} \twoheadrightarrow \mathsf{N}$ is *relative to* $\mathcal{F}$ when no redex underlined by a name $\mathsf{R} \notin \mathcal{F}$ is reduced along the marked reduction corresponding to ρ. For instance, the reduction in Example 5.1.1 is relative to $\{\mathsf{R}, \mathsf{S}\}$.

A reduction $\rho : \mathsf{M} \twoheadrightarrow \mathsf{N}$ is a (complete) *development* of a set of redexes $\mathcal{F}$ of M when ρ is relative to $\mathcal{F}$ and, after the marked reduction corresponding to ρ, N does not contain any redex underlined by a name $\mathsf{R} \in \mathcal{F}$.

Theorem 5.1.3 (finite developments) *Let $\mathcal{F}$ be a set of redexes of a λ-term.*

(i) *There is no infinite reduction relative to $\mathcal{F}$.*

(ii) *All developments end at the same term.*

(iii) *For any redex R of M and any pair of developments ρ and σ relative to $\mathcal{F}$, we have that $\mathsf{R}/\rho = \mathsf{R}/\sigma$.*

A first consequence of the previous lemma is that there is no ambiguity in writing $\mathcal{F} : \mathsf{M} \to \mathsf{N}$, for we may assume that this is a notation for all the developments of $\mathcal{F}$. Besides, we get in this way a parallel reduction of λ-terms in which a whole set of redexes is simultaneously reduced at each step. Thus, the composition $\mathcal{F}_1\mathcal{F}_2 \cdots \mathcal{F}_k$ denotes a sequence of parallel reduction steps $\mathcal{F}_i : \mathsf{M}_{i-1} \to \mathsf{M}_i$. In particular, as there is no restriction on the shape of the previous sets, $\mathcal{F}_i$ might even be empty; in which case $\mathsf{M}_{i-i} = \mathsf{M}_i$. In spite of this, we stress that $\emptyset \neq \epsilon$ (ϵ being the empty reduction), then in the equivalence by permutation that we will define later, we will have to explicitly state that the empty reduction and the reduction relative to an empty set of redexes are equivalent.

The notion of residual can be directly extended to parallel reductions, *i.e.*, $\mathsf{R}/\mathcal{F}$ is the residual of the redex R under (any development of) $\mathcal{F}$ (let us note that by Theorem 5.1.3 such a definition is not ambiguous). Furthermore, residuals can be also extended to a set of redexes, taking $\mathsf{R}' \in \mathcal{G}/\mathcal{F}$ iff $\mathsf{R}' \in \mathsf{R}/\mathcal{F}$ for some $\mathsf{R} \in \mathcal{G}$.

Of particular relevance is the reduction $\mathcal{F} \sqcup \mathcal{G} = \mathcal{F}(\mathcal{G}/\mathcal{F})$, that is, given two sets of redexes $\mathcal{F}$ and $\mathcal{G}$ of a λ-term, the reduction obtained reducing first the set $\mathcal{F}$ and then the residual of $\mathcal{G}$. In terms of developments, it is not difficult to realize that, appending a development σ of $\mathcal{G}/\mathcal{F}$ to a development ρ of $\mathcal{F}$, we get a development $\rho\sigma$ of $\mathcal{F} \cup \mathcal{G}$. Nevertheless, $\mathcal{F} \cup \mathcal{G}$ and $\mathcal{F} \sqcup \mathcal{G}$ differ, for they denote different sets of reduction sequences. Hence, assuming parallel reduction as atomic, we are not allowed to equate them. What we expect to equate instead are reductions in which redexes are permuted preserving their grouping, *i.e.*, as in the two reductions $\mathcal{F} \sqcup \mathcal{G}$ and $\mathcal{G} \sqcup \mathcal{F}$.

Lemma 5.1.4 (parallel moves) *Let $\mathcal{F}$ and $\mathcal{G}$ be set of redexes of a λ-term M. We have that:*

(i) *$\mathcal{F} \sqcup \mathcal{G}$ and $\mathcal{G} \sqcup \mathcal{F}$ ends at the same expression;*
(ii) *$\mathcal{H}/(\mathcal{F} \sqcup \mathcal{G}) = \mathcal{H}/(\mathcal{G} \sqcup \mathcal{F})$.*

The main consequence of the previous lemma is that the equivalence $\mathcal{F} \sqcup \mathcal{G} \equiv \mathcal{G} \sqcup \mathcal{F}$ is sound. We can thus use it as the core of the equivalence of reduction by permutation, the other equations being the ones induced by the fact that we want to obtain congruence with respect to composition of reductions.

Definition 5.1.5 (permutation equivalence) We define the *permutation equivalence* of reductions $\equiv$ as the smallest congruence with respect to composition of reductions satisfying the parallel moves lemma and elimination of empty steps. Namely, $\equiv$ is the smallest equivalence such that:

(i) $\mathcal{F} \sqcup \mathcal{G} \equiv \mathcal{G} \sqcup \mathcal{F}$, when $\mathcal{F}$ and $\mathcal{G}$ are set of redexes of the same λ-term;

(ii) $\emptyset = \epsilon$;

(iii) $\rho\sigma\tau \equiv \rho\sigma'\tau$, when $\sigma \equiv \sigma'$.

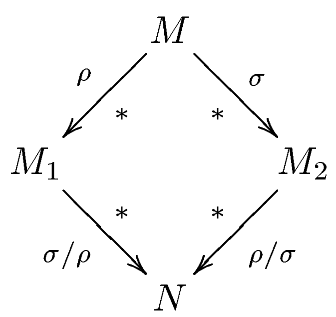
The notion of residual extends to reductions too. In fact, we have the following definition by induction on the length of the reduction:

$$\begin{aligned} \epsilon/\rho &= \epsilon \\ (\sigma\mathcal{F})/\rho &= (\sigma/\rho)(\mathcal{F}/(\rho/\sigma)) \end{aligned}$$

The previous definition allows us to conclude this part with the so-called diamond property.

Theorem 5.1.6 (diamond property) *For any pair of reductions* ρ *and* σ *starting from the same term, we have that* $\rho\,(\sigma/\rho) = \sigma\,(\rho/\sigma)$*.*

That can be graphically depicted by the following diamond:



Let us remark that the previous commuting diagram is indeed the base of the equivalence by permutation. In fact, two reductions ρ and σ are permutation equivalent if when composing instances of the previous diagram we can get a commuting diagram in which the composition of the external reductions yields ρ and σ.

The diamond property is a strong version of the so-called Church–Rosser (or confluence) property. In fact, it proves the confluence of the calculus, showing at the same time how to complete any pair of reductions ρ and σ in order to get the same result. Hence, an easy corollary of the previous result is the uniqueness of the normal form (if any).

5.1.2 Families of redexes

Definition 5.1.7 (copy) A redex S with history σ is a *copy* of a redex R with history ρ, written $\rho R \leq \sigma S$, if and only if there is a derivation τ such that $\rho\tau$ is permutation equivalent to σ ($\rho\tau \equiv \sigma$) and S is a residual of R with respect to τ ($S \in R/\tau$).

Definition 5.1.8 (family) The symmetric and transitive closure of the copy relation is called the *family* relation, and will be denoted with $\simeq$.

Explicitly, two redexes R and S with respective histories ρ and σ are in the same family ($\rho R \simeq \sigma S$) if and only if there is a finite sequence $\tau_0 T_0, \tau_1 T_1, \ldots, \tau_k T_k$, with $\tau_0 T_0 = \rho R$ and $\tau_k T_k \leq \sigma S$, such that either $\tau_{i-1} T_{i-1} \leq \tau_i T_i$ or $\tau_i T_i \leq \tau_{i-1} T_{i-1}$, for $i = 1, \ldots, k$. That, pictorially, gives rise to a sort of 'zig-zag' (see Figure 5.2).

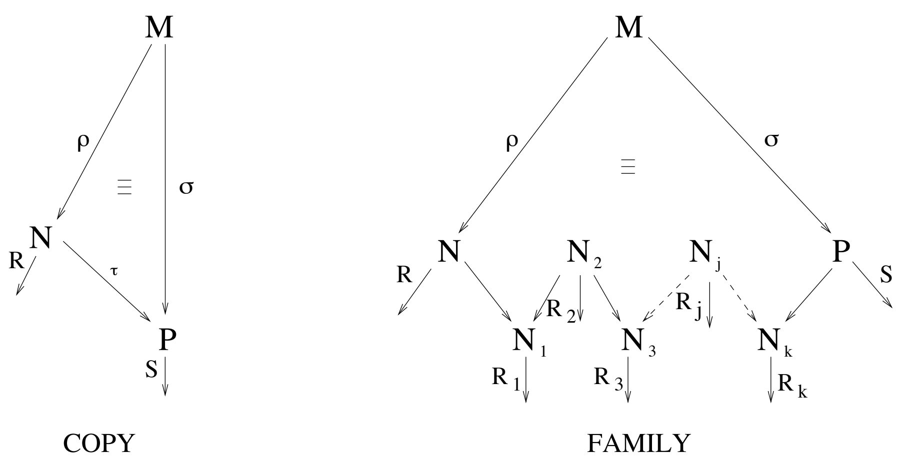


Fig. 5.2. Copy and family relations.

Example 5.1.9 Coming back to the example of Figure 5.1. We have:

$$RS_1T_1 \leq RS_1S_3T_3 \geq ST \leq RS_2S_4T_4 \geq RS_2T_2$$

From which we conclude that $RS_1T_1 \simeq RS_2T_2$.

Let us now prove a few properties of the copy and family relations.

Lemma 5.1.10 *Let* $\rho \equiv \rho'$ *and* $\sigma \equiv \sigma'$. *Then:*
(i) $\rho R \leq \sigma S$ *iff* $\rho' R \leq \sigma' S$
(ii) $\rho R \simeq \sigma S$ *iff* $\rho' R \simeq \sigma' S$

Proof Obvious, since $\equiv$ is a congruence for composition. □

Although the previous lemma looks straightforward, its relevance is not negligible. Indeed, it says that in order to check that two redexes are in the same family, we can consider any other history of the redexes, provided that they are permutation equivalent to the given ones. In particular, it allows us to restrict our analysis to *standard* derivations. In fact, let us recall that a derivation $R_1 R_2 \cdots$ is standard when R_i is not a residual of any redex at the left of R_j, for any i and any $j < i$. In the case of parallel reductions, the derivation $\mathcal{F}_1 \mathcal{F}_2 \cdots$ is standard when the previous proviso holds assuming that R_i and R_j are the leftmost-outermost redexes of the respective sets $\mathcal{F}_i$ and $\mathcal{F}_j$. A relevant result of λ-calculus—the so-called *standardization theorem*—ensures that, for any reduction, there exists an equivalent one which is standard.

The permutation equivalence allows us to extend the usual preorder given by the prefix relation between reduction sequences. Namely, let us define $\rho \sqsubseteq \sigma$, when $\rho\tau \equiv \sigma$ for some reduction τ.

Lemma 5.1.11 $\rho R \leq \sigma S$ *iff* $\rho \sqsubseteq \sigma$ *and* $S \in R/(\sigma/\rho)$.

Proof (⇒) By definition, if $\rho R \leq \sigma S$ there exists τ such that $\rho\tau \equiv \sigma$ and $S \in R/\tau$. Thus, $\rho \sqsubseteq \sigma$, by definition of $\sqsubseteq$. Moreover, $\tau \equiv \sigma/\rho$, which implies $R/\tau = R/(\sigma/\rho)$ and $S \in R/(\sigma/\rho)$. (⇐) Just take $\tau = (\sigma/\rho)$. □

Moreover, because of the following property of interpolation:

Lemma 5.1.12 (interpolation) *For any* $\rho \sqsubseteq \sigma \sqsubseteq \tau$ *and* $\rho R \leq \tau T$, *there exists a redex* S *such that* $\rho R \leq \sigma S \leq \tau T$.

Proof By assumption, there exist reductions ρ', σ' and τ' such that $\rho\rho' \equiv \sigma$, $\sigma\sigma' \equiv \tau$, $\rho\tau' \equiv \tau$ and $T \in R/\tau'$. Thus $\rho\tau' \equiv \rho\rho'\sigma'$, and by left-cancellation $\tau' \equiv \rho'\sigma'$. Since R has a residual T after $\tau' \equiv \rho'\sigma'$, then it must also have a residual S after ρ'. □

We can see that the copy relation is decidable, since for any σS, the minimal redex ρR such that $\rho R \leq \sigma S$ is unique.

Lemma 5.1.13 (uniqueness) *If* $\rho R_i \leq \sigma S$, *for* $i = 1, 2$, *then* $R_1 = R_2$.

Proof Since each redex is a residual of at most one ancestor. □

Exercise 5.1.14

(i) Prove that $\leq$ is a preorder.

(ii) Prove that $R \simeq \sigma S$ if and only if $S \in R/\sigma$. (*Hint*: For the if part, apply an induction on the definition of $\simeq$, using interpolation and uniqueness.)

5.2 Extraction

Intuitively, two redexes ρR and σS are in a same family if and only if they have been created 'in the same way' along σ and ρ. Nevertheless, this intuition is not easy to formalize, since 'creation' is a very complex operation in the λ-calculus.

A way to get rid of this complexity is to modify the calculus associating a label to each (sub)term. In this labeled version of the calculus, labels would be a trace of the history of each subterm, and in particular of the way in which each redex has been created. This technique, which generalizes an idea of Vuillemin for recursive program schemes, leads to the *labeled λ-calculus* that will be presented in section 5.3. Here, we give an alternative approach to zig-zag that does not involve any modification of the calculus.

Let us assume a simplification process (extraction) such that any redex ρR throws away all the redexes in its history ρ that are not relevant to the 'creation' of R. At the end of this process, we essentially obtain the 'causal history' of R (with respect to ρ). The causal histories of redexes could then be used to decide when two redexes are in the same family. Namely, we could say that ρR and σS are in the same family when the extraction process contracts them to the same redex τT. However, in order to achieve such a strong result, we immediately see that we have to fix some technical details. In particular, when several redexes participate in the creation of R and S, the corresponding casual histories might differ for the order in which such redexes are applied. Unfortunately, this would immediately lead back to permutation equivalence and zig-zag. Thus, as we want a unique 'linear' representation of equivalent causal histories (namely, a unique derivation), we are forced to organize redexes in some fixed order. But in view of Lemma 5.1.10, this does not seem a big problem, for restricting to *standard* derivations fulfils the uniqueness requirement and does not force any limitation.

In order to formally define the extraction relation, we need a few preliminary definitions. In the next subsection we will show that extraction gives indeed a decision procedure for the family relation. In fact, although the results of section 5.2.1 hold for redexes whose history is in standard form, this does not impact upon the decidability of zig-zag, for there is a recursive algorithm transforming a given reduction ρ into a (unique) standard reduction ρ_s such that $\rho \equiv \rho_s$.

We say that two derivations $\sigma : M \twoheadrightarrow N$ and $\rho : M \twoheadrightarrow P$ are *disjoint* if they contract redexes into disjoint subexpressions of M. This means that $\sigma : M = C[Q_1, Q_2] \twoheadrightarrow C[Q_1', Q_2] = N$, and $\rho : M = C[Q_1, Q_2] \twoheadrightarrow C[Q_1, Q_2'] = P$, for some context $C[\bullet, \bullet]$ with two disjoint holes. We also recall that, in a redex $R = (\lambda x.M\ N)$, the subterm $\lambda x.M$ is the *function part* of R, while M is its *argument part.*

Finally, let us recall that σ/ρ (the residual of a derivation σ with respect to a derivation ρ) is defined inductively by:

$$\sigma/\rho = \begin{cases} \epsilon & \text{when } \sigma = \epsilon \\ (\sigma'/\rho)\,(R/(\rho/\sigma')) & \text{when } \sigma = \sigma' R \end{cases}$$

where ϵ is the empty reduction.

The next definition gives the only issue that poses some technical difficulties in the definition of extraction. The idea is that, given a redex $R = (\lambda x.\,M\ N)$ and a reduction $R\rho_i$ such that ρ_i works inside the ith instance N_i of the argument part N of R. Any redex created by $R\rho_i$ could have been created by the direct execution in N of the reduction ρ isomorphic to ρ_i. In order to formalize this point, let us introduce the notation $M^x[\bullet, \bullet, \ldots, \bullet]$ to represent the context obtained by replacing a hole for all the occurrences of a given variable x occurring free in M.

Definition 5.2.1 (parallelization) Let $R = (\lambda x.M\ N)$. Let $R\sigma$ be a derivation such that σ is internal to the ith instance of N in the contractum $M[N/x]$ of R (see the figure in item 4 of Definition 5.2.2). The reduction $\sigma \| R$ (σ *parallelized* by R), is inductively defined as follows:

$$\begin{array}{rcll} \epsilon \,\|\, R & = & \epsilon & \\ (S\sigma) \,\|\, R & = & (S'/R)\,((\sigma/\mathcal{F})\|(R/S')) & \quad \text{where } S \in S'/R,\ \mathcal{F} = S'/(RS) \end{array}$$

Intuitively, for σ corresponds to a reduction applied to the instance of N inserted in the ith hole of $M^x[\bullet, \ldots, \bullet]$, the reduction σ works on a subterm isomorphic to N. Hence, there exists a reduction

$$\sigma' : C[(\lambda x.\,M\ N)] \twoheadrightarrow C[(\lambda x.\,M\ N')]$$

internal to N and isomorphic to σ, such that

$$\sigma : C[M^x[N, \ldots, N, \ldots, N]] \twoheadrightarrow C[M^x[N, \ldots, N', \ldots, N]]$$

The order of R and σ could then be commuted obtaining $\sigma' \sqcup R$. Nevertheless, it is immediate that after $\sigma' \sqcup R$ all the instances of N in

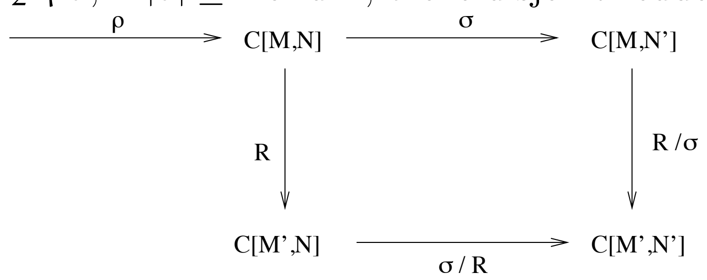
$C[M^x[N, \ldots, N, \ldots, N]]$ are contracted to M'. The reduction $\sigma' \sqcup R$ is indeed the parallelization of σ by R, that is, $\sigma'/R = \sigma' \sqcup R$ (see again item 4 of Definition 5.2.2).

In order to clarify how Definition 5.2.1 fits such an idea of parallelization, we remark that, by induction on $S\sigma$: *(i)* S' is internal to N; *(ii)* $\mathcal{F}$ is disjoint from σ; *(iii)* R/S' is a singleton $\{R'\}$; *(iv)* $\sigma/\mathcal{F}$ is internal to the ith instance of the argument of R' in its contractum. By which we conclude the soundness of Definition 5.2.1 and that, as we anticipated, $(S\sigma)\|R = (S'\sigma')/R$, for a suitable reduction σ' internal to N.

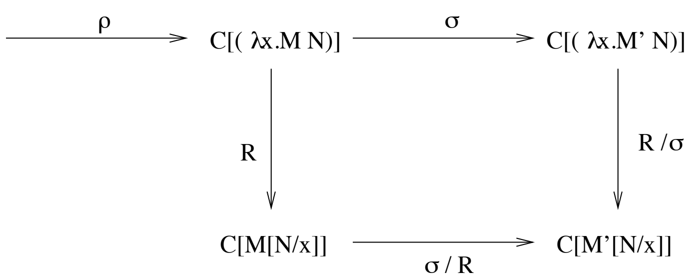
Definition 5.2.2 (extraction) The contraction by extraction $\rhd$ is the union of the following four relations:

(i) $\rho R S \; \rhd_1 \; \rho S'$, if $S \in S'/R$;

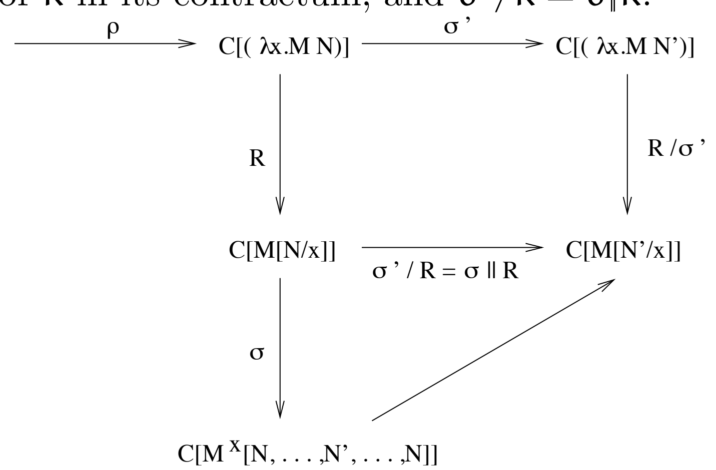
(ii) $\rho\,(R \sqcup \sigma) \; \rhd_2 \; \rho\sigma$, if $|\sigma| \geq 1$ and R, σ are disjoint reductions;



(iii) $\rho(R \sqcup \sigma) \; \rhd_3 \; \rho\sigma$, if $|\sigma| \geq 1$ and σ is internal to the function part of R;



(iv) $\rho R \sigma \; \rhd_4^i \; \rho\sigma'$, if $|\sigma| \geq 1$, σ is internal to the ith instance of the argument of R in its contractum, and $\sigma'/R = \sigma\|R$.



The *extraction* relation $\trianglerighteq$ is the transitive and reflexive closure of $\rhd$.

Example 5.2.3 Let us consider again the example in Figure 5.1. By parallelization, we have $RS_1T_1 \rhd ST$ and $RS_2T_2 \rhd ST$. By the second rule of the extraction relation, we have $RS_1S_3T_3 \rhd RS_1T_1$ and $RS_2S_4T_4 \rhd RS_2T_2$. Summarizing,

$$RS_1S_3T_3 \trianglerighteq ST \trianglelefteq RS_2S_4T_4$$

Which is the same result that we could have obtained by applying zig-zag (see Example 5.1.9).

Theorem 5.2.4 ([Lév80]) $\trianglerighteq$ *is confluent and strongly normalizing.*

Proof See [Lév80] or solve Exercise 5.2.6. □

Exercise 5.2.5 Let $\rho_i R_i$, $\sigma_i S_i$, and $\tau_i T_i$ be the redexes with history relative to the example in Figure 5.1. Applying $\trianglerighteq$ to each of them, verify that extraction is confluent and strongly normalizing. Furthermore, check that: (*i*) R is the unique normal form of any $\rho_i R_i$; (*ii*) S is the unique normal form of any $\sigma_i S_i$; (*iii*) ST is the unique normal form of any $\tau_i T_i$.

Exercise 5.2.6 Prove the following fact:

> Let R and S be two distinct redexes in a term M. If T, T_1, T_2 are such that $T \in T_1/(R/S)$, $T \in T_2/(S/R)$, then there exists some redex T' such that $T \in T'/(R \sqcup S)$ and $T \in T'/(S \sqcup R)$.

Use the previous result to prove Theorem 5.2.4.

5.2.1 Extraction and families

In this section, we shall consider standard derivations only.

Proposition 5.2.7 *If there exists a reduction* τT *such that* $\rho R \trianglerighteq \tau T \trianglelefteq \sigma S$*, then* $\rho R \simeq \sigma S$*.*

Proof It is enough to observe that, if $\rho R \trianglerighteq \tau T$, then $\rho R \simeq \tau T$. In the first three cases in the definition of $\rhd$, we obviously have $\tau T \leq \rho R$. In the last case (see Figure 5.3) we have $\rho = \rho' R' \rho''$ and $\tau = \rho' \tau'$, for some R'. Let $\nu = \tau'/(R'\rho'')$. Then $\rho''\nu \equiv \tau'/R'$. Moreover, R has a unique residual T' after ν, and this must also be a residual of T after R'/τ'. (Note that T is internal to N' in $C[(\lambda.x\ N')]$, and that R is the image of T in $C[M^x[N, \dots, N', \dots, N]]$.) Thus, $\rho R \leq \rho'(R' \sqcup \tau')T' \geq \tau T$, which implies $\rho R \simeq \tau T$. □

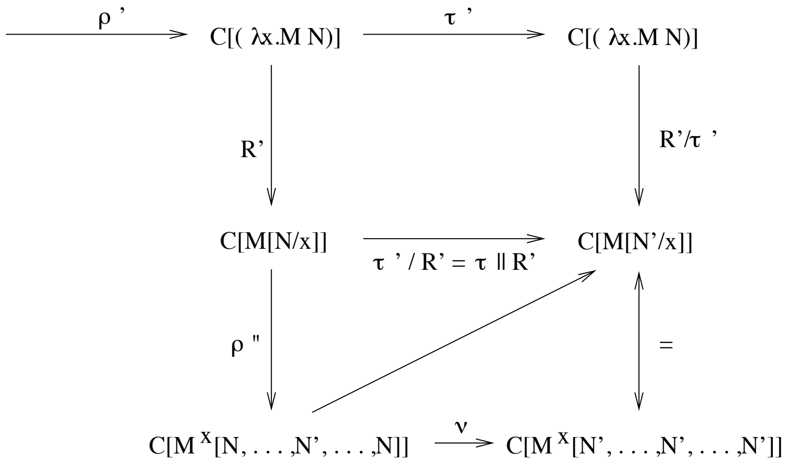


Fig. 5.3.

Also the converse of the previous proposition is true. But to prove it we need some preliminary lemmas.

Lemma 5.2.8 *If* $R \leq \sigma S$*, then* $\sigma S \trianglerighteq R$*.*

Proof If $R \leq \sigma S$, then $S \in R/\sigma$ and $\sigma S \trianglerighteq R$ by $\triangleright_1$. □

Lemma 5.2.9 *Let* $\rho R \leq \sigma S$*. There is a reduction* τT *such that* $\rho R \trianglerighteq \tau T \trianglelefteq \sigma S$*.*

Proof First of all, let us note that $\rho \sqsubseteq \sigma$, for by hypothesis $\rho R \leq \sigma S$. The proof continues then by induction on $|\sigma|$. The base case is immediate. In fact, when $|\sigma| = 0$, also $|\rho| = 0$, for $\rho \sqsubseteq \sigma$. The result follows then by Lemma 5.2.8. So, let us proceed with the induction case.
We can distinguish two subcases, according to the length of ρ. The easy one is $|\rho| = 0$, for the result follows again by Lemma 5.2.8. Thus, let $\sigma = S'\sigma'$ and $\rho = R'\rho'$. If $S' = R'$, then $\rho' R \leq \sigma' S$ and, by the induction hypothesis, there exists $\tau' T$ such that $\rho' R \trianglerighteq \tau' T \trianglelefteq \sigma' S$. Then, for $\tau = R'\tau'$, $\rho R \trianglerighteq \tau T \trianglelefteq \sigma S$.
Summarizing, we have left to prove the case $\sigma = S'\sigma'$, $\rho = R'\rho'$, and $R' \neq S'$. The redex R' cannot be external or to the left of S', for otherwise $R'/\sigma' \neq \emptyset$, contradicting the hypothesis $\rho \sqsubseteq \sigma$. Since ρ is standard (recall the assumption at the beginning of this section), it can be decomposed as $\rho = \rho_f \sqcup \rho_a \sqcup \rho_d$ such that ρ_f, ρ_a and ρ_d are respectively internal to the function part of S', internal to the argument part of S', and disjoint from S' (to its right). Moreover, S' has a *unique* residual S'' after ρ (recall that S' is leftmost w.r.t. Rq'). We proceed

then by case analysis, according to the mutual positions of R and S'' in the final term of ρ.

(i) R is external or to the left of S'': There is a redex T external or to the left of S' such that $R \in T/\rho$ (easy induction on $|\sigma|$). Thus, $\rho R \trianglerighteq T$. That, by Proposition 5.2.7, implies $T \simeq \rho R$. Moreover, $T \simeq \sigma S$, for $\rho R \leq \sigma S$. Hence, $S \in T/\sigma$ (see Exercise 5.1.14) and $\sigma S \trianglerighteq T$ (by definition of $\triangleright_1$).

(ii) R is internal to the function part of S'': Let $\nu = (\rho_a \sqcup \rho_d)/\rho_f$. If S_f is the unique residual of S' after ρ_f, then $S_f/\nu = \{S''\}$. Moreover, ν is disjoint from the function part of S_f. Thus, there exists a redex R_f in the function part of S_f such that $R_f/\nu = \{R\}$, which implies both $\rho R \trianglerighteq \rho_f R_f$ and $\rho R \geq \rho_f R_f$. By transitivity, since $\rho R \leq \sigma S$, we also have $\rho_f R_f \leq \sigma S$. By Lemma 5.1.11, $S \in R_f/(\sigma/\rho_f)$. Then, let $\rho'_f = \rho_f/S'$ and $R'_f = R_f/S_f$ (*i.e.*, $(\rho_f R_f)/S' = \rho'_f R'_f$). We have that $\rho'_f R'_f \leq \sigma' S$ and $S'\rho'_f R'_f \trianglerighteq \rho_f R_f$ (see Figure 5.4). Since ρ_f is in the function part of S' and ρ_f

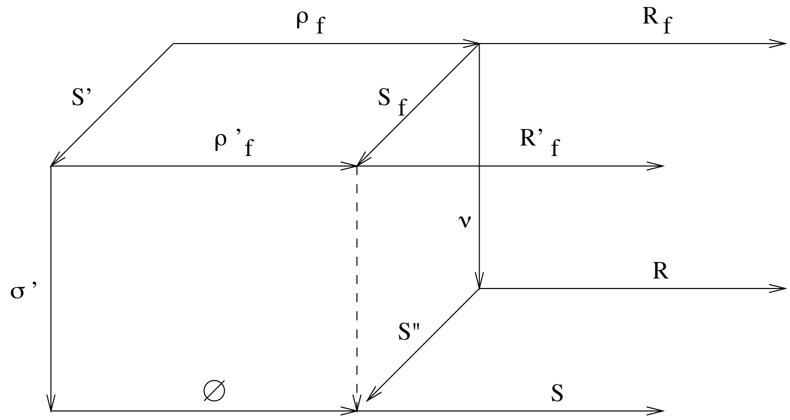


Fig. 5.4.

is standard, $S'\rho'_f$ and then ρ'_f are standard too. We can now apply the induction hypothesis to $\rho'_f R'_f \leq \sigma' S$, concluding that there is $\tau' T'$ such that $\rho'_f R'_f \trianglerighteq \tau' T' \trianglelefteq \sigma' S$. Therefore, $S'\rho'_f R'_f \trianglerighteq S'\tau' T' \trianglelefteq \sigma S$ (let us note that also τ' is in the function part of S', and then that $S'\tau'$ too is standard). But we have already seen that $S'\rho'_f R'_f \trianglerighteq \rho_f R_f$. So, by the Church–Rosser property of $\trianglerighteq$, there exists a derivation τT such that $S'\tau' T' \trianglerighteq \tau T \trianglelefteq \rho_f R_f$. Thus, $\sigma S \trianglerighteq \tau T \trianglelefteq \rho R$.

(iii) R is disjoint from S'' or to its right: There is again some R_d disjoint from the residual S_d of S' after ρ_d, such that $\rho R \trianglerighteq \rho_d R_d$ and $\rho R \geq \rho_d R_d$. Then we proceed as in the previous case.

(iv) R is in the argument part of S'': There is a redex R_a in the argument part of the residual S_a of S' after ρ_a, such that $\rho R \trianglerighteq \rho_a R_a$ and $\rho R \geq \rho_a R_a$. Unfortunately, the previous reasoning does not follow so simply in this case, since $\rho''_a = \rho_a/S'$ might not be standard. In fact, let us note that ρ''_a is the union of disjoint reductions, each of them internal to a different instance of the argument of S_a in its contractum. Anyhow, it is still true that $\rho_a R_a \leq \sigma S$, for $\rho R \leq \sigma S$. Thus, $S \in R_a/(\sigma/\rho_a)$ and there is a redex $R''_a \in R_a/S_a$ such that $\rho''_a R''_a \leq \sigma' S$. Moreover, R''_a is internal to some instance, say the ith one, of the argument part of S_a in its contractum. Let us take the component ρ'_a of ρ''_a internal to such an instance. There is a redex R'_a such that $\rho'_a R'_a \leq \rho R$, with $\rho'_a R'_a$ standard. Moreover, $S' \rho'_a R'_a \trianglerighteq \rho_a R_a$, for $\rho_a R_a/S' = (\rho'_a R'_a) \| S'$ by $\triangleright_4$; and $\rho'_a R'_a \leq \sigma' S$, for $\rho''_a R''_a \leq \sigma' S$. We can now proceed as in case 2. By induction hypothesis, there is $\tau' T'$ such that $\rho'_a R'_a \trianglerighteq \tau' T' \trianglelefteq \sigma' S$. Therefore, $S' \rho'_a R'_a \trianglerighteq S' \tau' T' \trianglelefteq \sigma S$. But since $S' \rho'_a R'_a \trianglerighteq \rho_a R_a$, by the Church–Rosser property of $\trianglerighteq$, there exists τT such that $S' \tau' T' \trianglerighteq \tau T \trianglelefteq \rho_a R_a$. Thus, $\sigma S \trianglerighteq \tau T \trianglelefteq \rho R$.

□

Let us comment on the structure of the previous proof. It follows exactly the definition of $\trianglerighteq$. In fact, each subcase in Definition 5.2.2 yields a corresponding subcase in the non-trivial part of the proof. It is indeed true that this is the actual reason for the four cases in the definition of extraction.

Proposition 5.2.10 *If* $\rho R \simeq \sigma S$, *then* $\rho R \trianglerighteq \tau T \trianglelefteq \sigma S$ *for some* τT.

Proof By definition of family, $\rho R \simeq \sigma S$ if and only if there exists a chain of $\rho_i R_i$ such that $\rho_0 R_0 = \rho R$, $\rho_n R_n = \sigma S$, and for all $1 \leq i \leq n$ either $\rho_{i-1} R_{i-1} \leq \rho_i R_i$ or $\rho_i R_i \leq \rho_{i-1} R_{i-1}$ (recall that we can assume without loss of generality that all the ρ_i are standard). By Lemma 5.2.9, there exists $\tau_i T_i$ such that $\rho_{i-1} R_{i-1} \trianglerighteq \tau_i T_i \trianglelefteq \rho_i R_i$, for every i. By the confluence of $\trianglerighteq$, we conclude then that there exists τT such that $\rho R \trianglerighteq \tau T \trianglelefteq \sigma S$. □

Theorem 5.2.11 (decidability of extraction) *Let* σ *and* ρ *be two* standard *reductions. Then* $\rho R \simeq \sigma S$ *if and only if* $\rho R \trianglerighteq \tau T \trianglelefteq \sigma S$ *for some* τT.

Proof By Proposition 5.2.7 and Proposition 5.2.10. □

The previous result not only establishes a full correspondence between zig-zag and extraction, but also gives an *effective* procedure for deciding the family relation. In fact, given two redexes ρR and σS, there is an effective way for deriving the standard reductions ρ_s and σ_s respectively equivalent to ρ and σ by permutation. Then, as extraction is effective and terminating, we can compute the $\trianglerighteq$ canonical forms $\rho'_s R'$ and $\sigma'_s S'$ of $\rho_s R$ and $\sigma_S S$. If and only if $\rho'_s R' = \sigma'_s S'$ the redexes ρR and σS are in the same family. The previous considerations can be summarized as follows.

Definition 5.2.12 (canonical derivation) Every standard derivation σu in normal form with respect to $\trianglerighteq$ will be called *canonical*.

Corollary 5.2.13 (canonical representative) *The* canonical representative *of a family can be effectively derived from each member* ρR *of the family: it is the unique canonical derivation* $\rho_c R_c$ *such that* $\rho_s R \trianglerighteq \rho_c R_c$*, where* ρ_s *is the standard derivation equivalent to* ρ*. Each family of redexes has a unique canonical representative.*

5.3 Labeling

The *labeled* λ*-calculus* is an extension of λ-calculus proposed by Lévy in [Lév78].

Let $L = \{a, b, \ldots\}$ be a denumerable set of *atomic labels*. The set $\mathbf{L}$ of *labels*, ranged over by $\alpha, \beta, \ldots$, is defined as the set of words over the alphabet L, with an arbitrary level of nested underlinings and overlinings.

Formally, $\mathbf{L}$ is the smallest set containing L and closed with respect to the following formation rules:

(i) if $\alpha \in \mathbf{L}$ and $\beta \in \mathbf{L}$, then $\alpha\beta \in \mathbf{L}$;
(ii) if $\alpha \in \mathbf{L}$, then $\underline{\alpha} \in \mathbf{L}$;
(iii) if $\alpha \in \mathbf{L}$, then $\overline{\alpha} \in \mathbf{L}$.

The operation of concatenation $\alpha\beta$ is supposed to be associative.

The set $\Lambda_V^{\mathbf{L}}$ of labeled λ-terms over a set V of variables and a set $\mathbf{L}$ of labels is defined as the smallest set containing:

(i) x^α, for any $x \in V$ and $\alpha \in \mathbf{L}$;
(ii) $(M\ N)^\alpha$, for all $M, N \in \Lambda_V^{\mathbf{L}}$ and $\alpha \in \mathbf{L}$;

(iii) $(\lambda x.M)^{\alpha}$, for any $M \in \Lambda_V^{\mathbf{L}}$ and $\alpha \in \mathbf{L}$.

As usual, we shall identify terms up to α-conversion.

The concatenation $\alpha \cdot M$ of a label α with a labeled term M is defined as follows:

(i) $\alpha \cdot x^{\beta} = x^{\alpha\beta}$;
(ii) $\alpha \cdot (M\ N)^{\beta} = (M\ N)^{\alpha\beta}$;
(iii) $\alpha \cdot (\lambda x.M)^{\beta} = (\lambda x.M)^{\alpha\beta}$.

The substitution $M[N/x]$ of a free variable x for a labeled λ-term N in a labeled λ-term M, is inductively defined by:

(i) $x^{\alpha}[N/x] = \alpha \cdot N$;
(ii) $y^{\alpha}[N/x] = y^{\alpha}$, when $y \neq x$;
(iii) $(M_1\ M_2)^{\alpha}[N/x] = (M_1[N/x]\ M_2[N/x])^{\alpha}$;
(iv) $(\lambda x.\, M)^{\alpha}[N/x] = (\lambda x.\, M)^{\alpha}$;
(v) $(\lambda y.\, M)^{\alpha}[N/x] = (\lambda y.\, M[N/x])^{\alpha}$.

In item (v) above, N is free for y in M, that is, no free variable of N is captured by the binder of y. This is not a limitation, due to our assumption on α-conversion, we can always suitably rename the variable y in $\lambda y.\, M$.

In the labeled system, β-reduction is defined by the following rule:

$$((\lambda x.M)^{\alpha}\ N)^{\beta} \to \beta \cdot \overline{\alpha} \cdot M[\underline{\alpha} \cdot N/x]$$

The *degree* of a redex $R = ((\lambda x.M)^{\alpha}\ N)^{\beta}$ is the label α of its function part (notation: $\mathsf{degree}(R). = \alpha$).

The formal presentation of the labeled λ-calculus given above should not scare the reader. The main idea is indeed very simple, once shifting from the concrete syntax of terms to their graphical representation as syntax trees. In this context, to add a label to each subterm corresponds to marking each edge in the tree by a label. The degree of a redex $R = ((\lambda x.M)^{\alpha}\ N)^{\beta}$ is the label of the edge between the corresponding @-node of $(\lambda x.M\ N)$ and the λ-node of $\lambda x.M$. Firing the redex, the degree of R is captured between the labels of the other edges incident to the nodes in R. Namely, between the label of the application of R and the label of the body of the abstraction in R, and between the label of any occurrence of the variable in R and the label of the argument part of R. In the contractum, the pairs of edges corresponding to the previous pairs of labels are replaced by new connections. Apart from lining, the

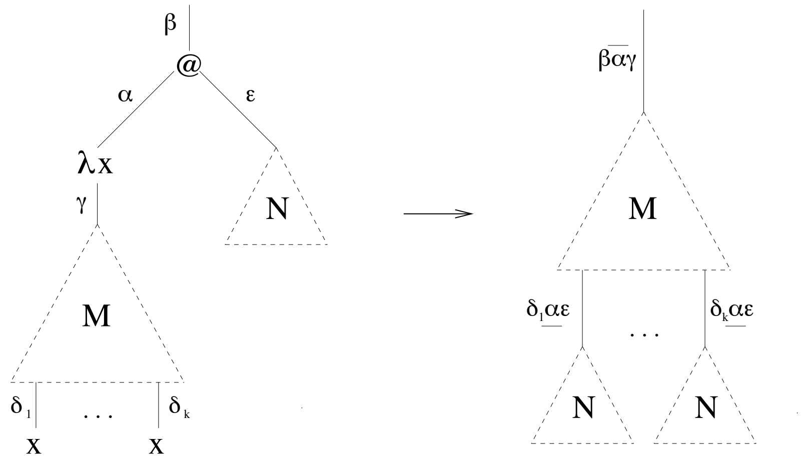


Fig. 5.5. Labeled β-reduction.

labels of these new edges are obtained by composing the labels of the corresponding edges in the natural way (see Figure 5.5).

Overlining and underlining respectively represent the two ways in which new connections are created by firing R: upwards (from the context to the body M of λx.M), and downwards (from the occurrences of the variable x in M to the instances of the argument N).

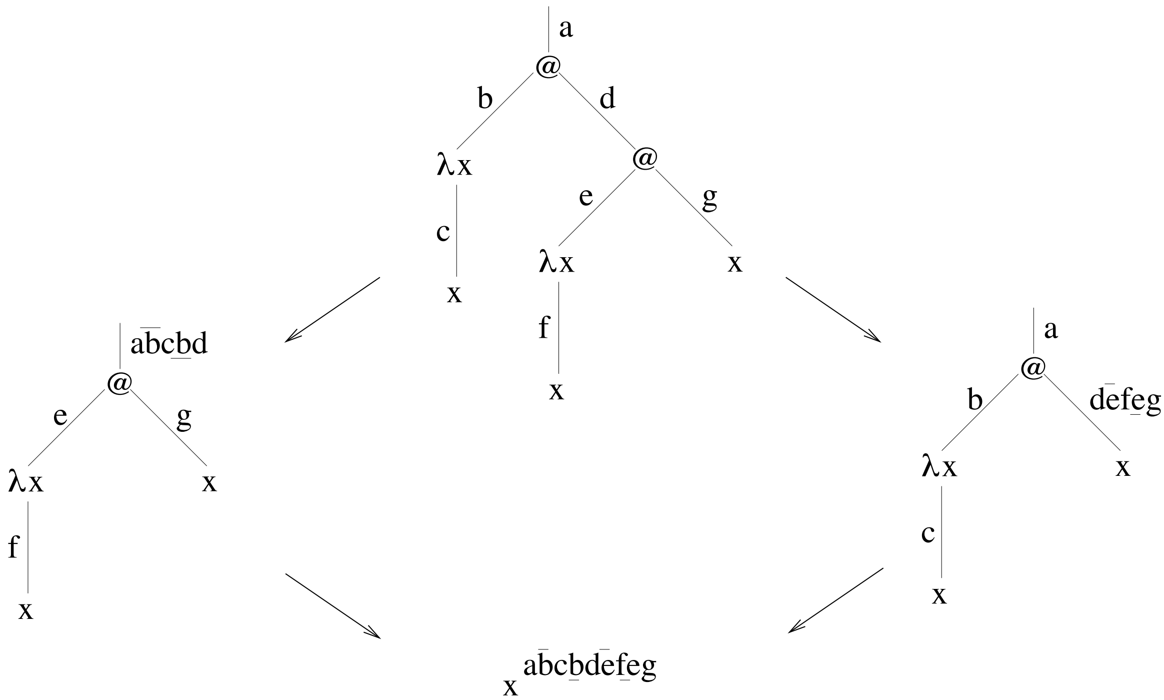


Fig. 5.6. (I (I x)).

Example 5.3.1 We already observed that, when contracting (I (I x)), after one step we get the same term (I x) whichever redex we reduced.

We pointed out that this equivalence is merely incidental. This situation is correctly handled by the labeled system, where labels allow one to distinguish between the way in which the two terms (I x) are obtained (see Figure 5.6). However, after one more step, the reduction ends with the same term also in the labeled system (Figure 5.6). In fact, labeling preserves confluence of the calculus, as we will show in the next section.

Example 5.3.2 In the λ-calculus, reducing (Δ Δ) we get a reduction sequence composed of an infinite number of copies of (Δ Δ). As for the previous example, in the labeled calculus this is no longer true. In fact, the reduction of (Δ Δ) gives rise to an infinite sequence of distinguished labeled terms (see Figure 5.7).

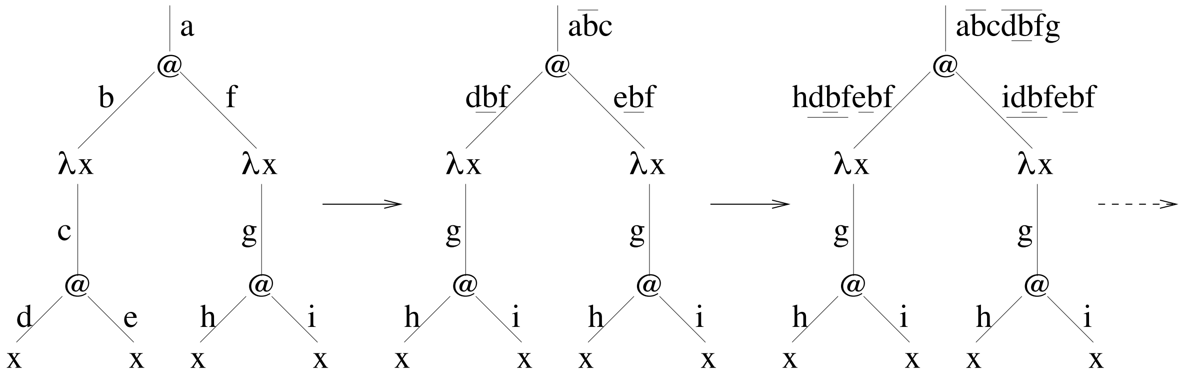


Fig. 5.7. (λx.(x x) λx.(x x)).

5.3.1 Confluence and standardization

In this section we shall prove the Church–Rosser and standardization properties for the labeled λ-calculus.

For the sake of the proof of such properties, we will consider a suitable extension of β-reduction. Namely, we will assume that the degrees of the redex contracted reducing terms verify a given predicate $\mathcal{P}$ on **L**. According to this, the reduction $((\lambda x.M)^{\alpha}\ N)^{\beta} \to \beta \cdot \overline{\alpha} \cdot M[\underline{\alpha} \cdot N/x]$ will be considered legal only when $\mathcal{P}(\alpha)$ is true.

The introduction of this predicate does not cause any loss of expressiveness, as usual β-reduction corresponds to the case in which $\mathcal{P}$ is always true. Furthermore, the use of $\mathcal{P}$ allows one to recover in a very simple way many other kinds of labelings considered in the literature, without a sensible complication of the theory of the calculus.

The schema of the proofs of confluence and standardization is the following:

(i) We shall start proving that labeled λ-calculus is locally confluent, for any choice of the predicate $\mathcal{P}$.
(ii) We shall directly prove standardization for strongly normalizable terms.
(iii) Since it is well known that local confluence implies confluence for strongly normalizable terms, the goal will be to exploit the previous results in conjunction with some suitable predicate $\mathcal{P}$ ensuring strong normalization. To this purpose, we shall give some sufficient conditions for $\mathcal{P}$ that imply strong normalization of every term (taking into account reductions that are legal with respect to $\mathcal{P}$ only).
(iv) Finally, we shall observe that, for any pair of reductions $\rho : M \twoheadrightarrow N$ and $\rho' : M \twoheadrightarrow N'$, we can construct a predicate $\mathcal{P}$ satisfying the above mentioned sufficient conditions, for which the reductions ρ and ρ' are legal.

The proof of local confluence is preceded by some lemmas used to prove that substitution behaves well with respect to labels. Namely, that the usual property of λ-calculus $M[N/x] \twoheadrightarrow M'[N'/x]$, when $M \twoheadrightarrow M'$ and $N \twoheadrightarrow N'$, holds in the labeled case too.

Lemma 5.3.3
(i) $\alpha \cdot (M[N/x]) = (\alpha \cdot M)[N/x]$;
(ii) $M[N/x][N'/y] = M[N'/x][N[N'/y]/x]$, *when* $x \neq y$ *and* x *does not occur free in* N'.

Proof An easy induction on the structure of M. □

Lemma 5.3.4 *If* $M \twoheadrightarrow M'$, *then* $M[N/x] \twoheadrightarrow M'[N/x]$.

Proof Let us prove first the case $M \to M'$ by structural induction on M:

(i) $M = x^\alpha$. This case is vacuous.
(ii) $M = (\lambda x.M_1)^\alpha$. Since the redex must be internal to M_1, $M' = (\lambda x.M_1')^\alpha$, with $M_1 \to M_1'$. By induction hypothesis, $M_1[N/x] \to M_1'[N/x]$, and thus $M[N/x] \to M'[N/x]$.
(iii) $M = (M_1\ M_2)^\alpha$, and the redex is internal to M_1 or M_2. This case is similar to the previous one.

(iv) $M = ((\lambda y.M_1)^\alpha\ M_2)^\beta$, $M' = \beta\overline{\alpha} \cdot M_1[\underline{\alpha} \cdot M_2/y]$ and $\mathcal{P}(\alpha)$ is true. By α-conversion we may suppose $x \neq y$ and y not free in N. Then, we have $M[N/x] = ((\lambda y.M_1[N/x])^\alpha\ M_2[N/x])^\beta$, and

$$\begin{aligned} M'[N/x] &= (\beta\overline{\alpha} \cdot M_1[\underline{\alpha} \cdot M_2/y])[N/x] \\ &= \beta\overline{\alpha} \cdot M_1[\underline{\alpha} \cdot M_2/y][N/x] \\ &\quad \text{by Lemma 5.3.3.(i)} \\ &= \beta\overline{\alpha} \cdot M_1[N/x][(\underline{\alpha} \cdot M_2)[N/x]/y] \\ &\quad \text{by Lemma 5.3.3.(ii)} \\ &= \beta\overline{\alpha} \cdot M_1[N/x][\underline{\alpha} \cdot M_2[N/x]/y] \\ &\quad \text{by Lemma 5.3.3.(i)} \end{aligned}$$

Since $\mathcal{P}(\alpha)$ is true, $M[N/x] \to M'[N/x]$.

By iteration, we get the case $M \twoheadrightarrow M'$. □

Lemma 5.3.5 *If* $M \twoheadrightarrow M'$, *then* $\alpha \cdot M \twoheadrightarrow \alpha \cdot M'$.

Proof Trivial. □

Lemma 5.3.6 *If* $N \twoheadrightarrow N'$, *then* $M[N/x] \twoheadrightarrow M[N'/x]$.

Proof By structural induction on M:

(i) $M = x^\alpha$. Then, $M[N/x] = \alpha \cdot N$, $M[N'/x] = \alpha \cdot N'$, and $\alpha \cdot N \twoheadrightarrow \alpha \cdot N'$, by Lemma 5.3.5.

(ii) $M = y^\alpha$, where $y \neq x$. Obvious.

(iii) $M = (\lambda x.M_1)^\alpha$. We have that $M[N/x] = (\lambda x.M_1[N/x])^\alpha$ (up to α-conversion) and that $M[N'/x] = (\lambda x.M_1[N'/x])^\alpha$. By induction hypothesis, $M_1[N/x] \twoheadrightarrow M_1[N'/x]$, and thus $M[N/x] \twoheadrightarrow M[N'/x]$.

(iv) $M = (M_1\ M_2)^\alpha$. Analogous to the previous case.

□

Corollary 5.3.7 *If* $M \twoheadrightarrow M'$ *and* $N \twoheadrightarrow N'$ *then,* $M[N/x] \twoheadrightarrow M'[N'/x]$.

Proof By Lemma 5.3.4 and Lemma 5.3.6. □

Proposition 5.3.8 (local confluence) *For any pair of redexes* $M \stackrel{R}{\to} M'$ *and* $M \stackrel{S}{\to} M''$, *there exist* N *and two reductions* ρ *and* σ *such that* $M' \stackrel{\sigma}{\twoheadrightarrow} N$ *and* $M'' \stackrel{\rho}{\to} N$.

Proof By structural induction on M:

(i) $M = x^\alpha$. This case is vacuous.

(ii) $M = (\lambda x.M_1)^\alpha$. Since R and S must be internal to M_1, $M' = (\lambda x.M_1')^\alpha$, $M'' = (\lambda x.M_1'')^\alpha$, $M_1 \to M_1'$, and $M_1 \to M_1''$. By induction hypothesis, there exists N_1 such that $M_1' \twoheadrightarrow N_1$ and $M_1'' \twoheadrightarrow N_1$. Taking $N = (\lambda x.N_1)^\alpha$ we have done.

(iii) $M = (M_1\ M_2)^\alpha$, and both redexes R and S are either internal to M_1 or internal to M_2. This case is similar to the previous one.

(iv) $M = (M_1\ M_2)^\alpha$, R is internal to M_1 and S is internal to M_2 (or vice versa). In this case, we close the diamond in one step, firing the unique residuals of R and S in M'' and M', respectively.

(v) $M = ((\lambda y.M_1)^\alpha\ M_2)^\beta \xrightarrow{R} M' = \beta\overline{\alpha} \cdot M_1[\underline{\alpha} \cdot M_2/y]$. We distinguish two subcases:

  (a) S is internal to M_1. Let M_1' be the reduct of M_1 by S. Then, $M'' = ((\lambda y.M_1')^\alpha\ M_2)^\beta \to N = \beta\overline{\alpha} \cdot M_1'[\underline{\alpha} \cdot M_2/y]$. Since $M_1 \to M_1'$, by Lemma 5.3.4 and Lemma 5.3.5, $M' = \beta\overline{\alpha} \cdot M_1[\underline{\alpha} \cdot M_2/y] \to N = \beta\overline{\alpha} \cdot M_1'[\underline{\alpha} \cdot M_2/y]$.

  (b) S is internal to M_2. Let M_2' be the reduct of M_2 by S. Then, $M'' = ((\lambda y.M_1)^\alpha\ M_2')^\beta \to N = \beta\overline{\alpha} \cdot M_1[\underline{\alpha} \cdot M_2'/y]$. Since $M_2 \to M_2'$, by Lemma 5.3.5 and Lemma 5.3.6, $M' = \beta\overline{\alpha} \cdot M_1[\underline{\alpha} \cdot M_2/y] \twoheadrightarrow N = \beta\overline{\alpha} \cdot M_1[\underline{\alpha} \cdot M_2'/y]$.

(vi) As in the previous case, inverting R and S.

It is immediate to check that, extending the definitions of residual and development to the labeled case (see Definition 5.1.2 and Theorem 5.1.3), we have indeed that $\rho = R/S$ and $\sigma = S/R$. □

Remark 5.3.9 Let us note again that the ρ and σ are legal for any predicate $\mathcal{P}$ such that both $\mathcal{P}(R)$ and $\mathcal{P}(S)$ are true. Indeed, since $\rho = R/S$ and $\sigma = S/R$, the degree of all the redexes fired in ρ (σ) is equal to $\mathsf{degree}(S)$ ($\mathsf{degree}(R)$).

It is a well known result of rewriting systems that local confluence implies confluence for strongly normalizable terms (in the literature, this property is known as the Newman Lemma). The proof is a simple induction on the length of the longest normalizing derivation for the term. So, we have the following corollary.

Proposition 5.3.10 (confluence) *Let M be a strongly normalizable term. If $\rho : M \twoheadrightarrow M'$ and $\sigma : M \twoheadrightarrow M''$, then there exist N such that $\rho' : M' \twoheadrightarrow N$ and $\sigma' : M'' \twoheadrightarrow N$, where $\rho' = \sigma/\rho$ and $\sigma' = \rho/\sigma$.*

Proof By Newman's Lemma and Lemma 5.3.8 we have confluence, that is, the existence of N, $\rho' : \mathsf{M}' \twoheadrightarrow \mathsf{N}$ and $\sigma' : \mathsf{M}'' \twoheadrightarrow \mathsf{N}$. Anyhow, since we want to prove that $\rho' = \sigma/\rho$ and $\sigma' = \rho/\sigma$, let us give the proof in full. Let us define a measure $\mathsf{depth}(\,.\,)$ on terms, such that $\mathsf{depth}(\mathsf{T})$ is the length of the longest reduction of T. The proof is by induction on $\mathsf{depth}(\mathsf{M})$.

(i) $\mathsf{depth}(\mathsf{M}) = 0$. M is in normal form. Hence, $\rho = \sigma = \epsilon$ and $\rho' = \sigma' = \epsilon$.

(ii) The cases $\rho = \epsilon$ or $\sigma = \epsilon$ are trivial, since we immediately see that $\rho' = \sigma$ and $\sigma' = \rho$. Hence, let us assume $\rho = \mathsf{M}' \xrightarrow{\mathsf{R}} \mathsf{P}' \overset{\rho_0}{\twoheadrightarrow} \mathsf{M}'$ and $\rho = \mathsf{M}' \xrightarrow{\mathsf{S}} \mathsf{P}'' \overset{\sigma_0}{\twoheadrightarrow} \mathsf{M}''$. By Lemma 5.3.8, there exists Q such that $\mathsf{S}/\mathsf{R} : \mathsf{T}' \twoheadrightarrow \mathsf{Q}$ and $\mathsf{R}/\mathsf{S} : \mathsf{T}'' \twoheadrightarrow \mathsf{Q}$. By definition, $\mathsf{depth}(\mathsf{T}'), \mathsf{depth}(\mathsf{T}'') < \mathsf{depth}(\mathsf{M})$. Hence, by induction hypothesis, there are Q' and Q'' such that $\rho_0/(\mathsf{S}/\mathsf{R}) : \mathsf{Q} \twoheadrightarrow \mathsf{Q}'$, $\sigma_0/(\mathsf{R}/\mathsf{S}) : \mathsf{Q} \twoheadrightarrow \mathsf{M}'$, $(\mathsf{S}/\mathsf{R})/\rho_0 : \mathsf{M}' \twoheadrightarrow \mathsf{Q}'$ and $(\mathsf{R}/\mathsf{S})/\sigma_0 : \mathsf{M}'' \twoheadrightarrow \mathsf{Q}''$. Finally, since $\mathsf{depth}(\mathsf{Q}) \leq \mathsf{depth}(\mathsf{T}')$ and $\mathsf{depth}(\mathsf{Q}) \leq \mathsf{depth}(\mathsf{T}'')$, by induction hypothesis we have that $(\sigma_0/(\mathsf{R}/\mathsf{S}))/(\rho_0/(\mathsf{S}/\mathsf{R})) : \mathsf{Q}' \twoheadrightarrow \mathsf{N}$ and $(\rho_0/(\mathsf{S}/\mathsf{R}))/(\sigma_0/(\mathsf{R}/\mathsf{S})) : \mathsf{Q}'' \twoheadrightarrow \mathsf{N}$, for some N. It is now an easy exercise to verify that

$$\mathsf{R}\rho_0((\mathsf{S}/\mathsf{R})/\rho_0)((\sigma_0/(\mathsf{R}/\mathsf{S}))/(\rho_0/(\mathsf{S}/\mathsf{R}))) = \sigma/\rho$$

and that

$$\mathsf{S}\sigma_0((\mathsf{R}/\mathsf{S})/\sigma_0)((\rho_0/(\mathsf{S}/\mathsf{R}))/(\sigma_0/(\mathsf{R}/\mathsf{S}))) = \sigma/\rho$$

□

Exercise 5.3.11 Verify the final two equivalences in the proof of Theorem 5.3.10. Moreover, we invite the reader to draw the reduction diagram corresponding to the previous proof.

Exercise 5.3.12 We invite the reader to reflect on why the strong normalization hypothesis is mandatory in Newman Lemma. Then, we invite the reader to prove that the following rewriting system

$$\mathtt{a} \rightarrow \mathtt{b} \qquad \mathtt{b} \rightarrow \mathtt{a} \qquad \mathtt{a} \rightarrow \mathtt{c} \qquad \mathtt{b} \rightarrow \mathtt{d}$$

(over the set of symbolsa $\{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}\}$) is not confluent, although it is locally confluent.

Let us now prove standardization for strongly normalizable terms.

The definition of standard and normal (leftmost-outermost) derivations are as usual:

- A reduction $M_0 \xrightarrow{R_1} M_1 \xrightarrow{R_2} \cdots \xrightarrow{R_k} M_k \xrightarrow{R_{k+1}} \cdots$ is *standard* when for any i, j such that $1 \leq i \leq j$, the redex R_j is not a residual of a redex in M_{i-1} to the left of R_i.
- A reduction $M_0 \xrightarrow{R_1} M_1 \xrightarrow{R_2} \cdots \xrightarrow{R_k} M_k \xrightarrow{R_{k+1}} \cdots$ is *normal* (or leftmost-outermost) when for any $i \geq 1$ the redex R_i is the leftmost redex in M_i.

In the following, we shall respectively use $M \overset{\mathsf{st}}{\twoheadrightarrow} N$ and $M \overset{\mathsf{norm}}{\twoheadrightarrow} N$ to denote standard and normal reductions.

It is immediate that any normal reduction is also standard and that normal reduction is uniquely determined. Namely, if $\rho : M \overset{\mathsf{norm}}{\twoheadrightarrow} N$ and $\rho' : M \overset{\mathsf{norm}}{\twoheadrightarrow} N'$, then ρ is a prefix of ρ', or vice versa.

Definition 5.3.13 The function Abs is defined as follows:

(i) $\mathsf{Abs}((\lambda x.M)^\alpha) = (\lambda x.M)^\alpha$;

(ii) $\mathsf{Abs}(M\ N)^\alpha) = \mathsf{Abs}(\alpha\overline{\beta} \cdot P[\underline{\beta} \cdot N/x])$ if $\mathsf{Abs}(M) = (\lambda x.P)^\alpha$.

So, $\mathsf{Abs}(M)$ is the first abstraction obtained by normal reduction of M. Obviously, Abs is not always defined.

Example 5.3.14 Variables are trivial examples of terms for which Abs is undefined. Nevertheless, for some terms the reason for which Abs is undefined are deeper. For instance, let us take again $(\Delta\,\Delta)$. Since, $(\Delta\,\Delta) \to (\Delta\,\Delta)$ is the only contraction of $(\Delta\,\Delta)$, we can never obtain an abstraction along its reduction.

Lemma 5.3.15 *Every standard reduction* $M \overset{\mathsf{st}}{\twoheadrightarrow} (\lambda x.N)^\alpha$ *can be decomposed in the following way:* $M \overset{\mathsf{norm}}{\twoheadrightarrow} \mathsf{Abs}(M) \twoheadrightarrow (\lambda x.N)^\alpha$

Proof Trivial. □

Proposition 5.3.16 *Let* M *be strongly normalizable. For any* $\rho : M \twoheadrightarrow N$*, there exists a corresponding standard reduction* $\sigma : M \overset{\mathsf{st}}{\twoheadrightarrow} N$*, such that* $\rho \equiv \sigma$*.*

Proof Let us call $\mathsf{depth}(M)$ the length of the longest normalizing derivation for M. The proof is by double induction over $\mathsf{depth}(M)$ and the structure of M.

If $\mathsf{depth}(M) = 0$ the result is trivial. If $\mathsf{depth}(M) > 0$, we proceed instead by structural induction on M:

(i) $M = x^{\alpha}$. This case is vacuous.

(ii) $M = (\lambda x.M_1)^{\alpha}$. Then, $N = (\lambda x.N_1)^{\alpha}$, and $M_1 \twoheadrightarrow N_1$. Furthermore, M_1 is a subterm of M and $\mathsf{depth}(M_1) = \mathsf{depth}(M)$. So, by induction hypothesis, $M_1 \overset{st}{\twoheadrightarrow} N_1$, and $M \overset{st}{\twoheadrightarrow} N$.

(iii) $M = (M_1\ M_2)^{\alpha}$. Let us distinguish two subcases:

(a) $N = (N_1\ N_2)^{\alpha}$, and the reduction $M \twoheadrightarrow N$ is composed of two separate reductions internal to M_1 and M_2. Then, by induction hypothesis, $M_1 \overset{st}{\twoheadrightarrow} N_1$, $M_2 \overset{st}{\twoheadrightarrow} N_2$, and $M = (M_1\ M_2)^{\alpha} \overset{st}{\twoheadrightarrow} (N_1\ M_2)^{\alpha} \overset{st}{\twoheadrightarrow} (N_1\ N_2)^{\alpha} = N$.

(b) The reduction $M \twoheadrightarrow N$ can be decomposed in the following way: $M = (M_1\ M_2)^{\alpha} \twoheadrightarrow ((\lambda x.N_1)^{\beta}\ M_2)^{\alpha} \to \alpha\overline{\beta} \cdot N_1[\underline{\beta} \cdot N_2/x] \twoheadrightarrow N$, where $M_2 \twoheadrightarrow N_2$ and $M_1 \twoheadrightarrow (\lambda x.N_1)^{\beta}$. By induction hypothesis, $M_1 \overset{st}{\twoheadrightarrow} (\lambda x.N_1)^{\beta}$, and by Lemma 5.3.15, $M_1 \overset{norm}{\twoheadrightarrow} (\lambda x.M_3) \overset{st}{\twoheadrightarrow} (\lambda x.N_1)^{\beta}$, where $\mathsf{Abs}(M_1) = (\lambda x.M_3)^{\beta}$. So, $M_3 \twoheadrightarrow N_1$. Moreover, $M_2 \twoheadrightarrow N_2$, and by Lemma 5.3.5 and Lemma 5.3.6, $\alpha\overline{\beta} \cdot M_3[\underline{\beta} \cdot M_2/x] \twoheadrightarrow \alpha\overline{\beta} \cdot N_1[\underline{\beta} \cdot N_2/x]$. Summing up,

$$\begin{aligned} M = (M_1\ M_2)^{\alpha} &\overset{norm}{\twoheadrightarrow} ((\lambda x.M_3)^{\beta}\ M_2)^{\alpha} \\ &\to \alpha\overline{\beta} \cdot M_3[\underline{\beta} \cdot M_2/x] \twoheadrightarrow \alpha\overline{\beta} \cdot N_1[\underline{\beta} \cdot N_2/x] \twoheadrightarrow N \end{aligned}$$

Let us now note that $\mathsf{depth}(\alpha\overline{\beta}\cdot M_3[\underline{\beta}\cdot M_2/x]) < \mathsf{depth}(M)$. Therefore, by induction hypothesis there exists a reduction $\alpha\overline{\beta} \cdot M_3[\underline{\beta} \cdot M_2/x] \overset{st}{\twoheadrightarrow} N$. So, we finally have the standard reduction:

$$\begin{aligned} M = (M_1\ M_2)^{\alpha} &\overset{norm}{\twoheadrightarrow} ((\lambda x.M_3)^{\beta}\ M_2)^{\alpha} \\ &\to \alpha\overline{\beta} \cdot M_3[\underline{\beta} \cdot M_2/x] \overset{st}{\twoheadrightarrow} N. \end{aligned}$$

We leave as an exercise for the reader verification that the reduction σ built in the proof and ρ are equivalent by permutation. □

Remark 5.3.17 The two propositions 5.3.10 and 5.3.16 hold for *every* choice of the predicate $\mathcal{P}$. In particular, assuming the usual notion of β-reduction, they hold for any subset of λ-terms that is strongly normalizable—for instance, strong normalization might be obtained by restricting our attention to terms typable according to a suitable typing discipline. Anyhow, we want to prove confluence and standardization for *any* λ-term. So, the strategy must be to restrict β-reduction, eliminating those reductions that cause non-termination. At the same time, we require that the reductions involved in the theorem remain legal. Hence, the next step is to find some sufficient conditions on $\mathcal{P}$ such that all terms of labeled λ-calculus become strongly normalizable.

Let us start by introducing some notation and definitions.

Notation 5.3.18 We shall use $\ell(M)$ to denote the external label of a labeled λ-term. In particular, $\ell(x^\alpha) = \ell((M\ N)^\alpha) = \ell((\lambda x.M)^\alpha) = \alpha$.

Definition 5.3.19 The height $h(\alpha)$ of a label α is the maximal nesting depth of overlinings and underlinings in it. Formally:

(i) $h(a) = 0$, when a is an atomic label;
(ii) $h(\alpha\beta) = \max\{h(\alpha), h(\beta)\}$;
(iii) $h(\overline{\alpha}) = h(\underline{\alpha}) = 1 + h(\alpha)$.

Definition 5.3.20 We say that the predicate $\mathcal{P}$ *has an upper bound*, if the set $\{h(\alpha) \mid \mathcal{P}(\alpha)\}$ has an upper bound.

Lemma 5.3.21 *If* $M \twoheadrightarrow M'$, *then* $h(\ell(M)) \leq h(\ell(M'))$.

Proof Obvious. □

Lemma 5.3.22 *Let* $T = (\ldots((M\ N_1)^{\beta_1}\ N_2)^{\beta_2} \ldots N_n)^{\beta_n}$. *If* $T \twoheadrightarrow (\lambda x.N)^\alpha$, *then* $h(\ell(M)) \leq h(\alpha)$.

Proof By induction on n. When $n = 0$, the result follows by the previous lemma. When $n > 0$, we must have:

- $(\ldots((M\ N_1)^{\beta_1}\ N_2)^{\beta_2} \ldots N_{n-1})^{\beta_{n-1}} \twoheadrightarrow (\lambda y.P)^\gamma$
- $N_n \twoheadrightarrow N'_n$
- $((\lambda y.P)^\gamma N'_n)^\beta \twoheadrightarrow \beta_n\overline{\gamma} \cdot P[\underline{\gamma} \cdot N'_n] \twoheadrightarrow (\lambda x.N)^\alpha$

So, by induction hypothesis,

$$h(\gamma) < h(\overline{\gamma}) \leq h(\ell(\beta_n\overline{\gamma} \cdot P[\underline{\gamma} \cdot N'_n])) \leq h(\alpha)$$

□

Lemma 5.3.23 *Any standard reduction* $M[N/x] \overset{st}{\twoheadrightarrow} (\lambda y.P)^\alpha$ *can be decomposed in one of the following two ways:*

(i) $M \twoheadrightarrow (\lambda y.M')^{\alpha}$ *and* $M'[N/x] \twoheadrightarrow P$.
(ii) $M \twoheadrightarrow M' = (\ldots((x\ M_1)^{\beta_1}\ M_2)^{\beta_2} \ldots M_n)^{\beta_n}$ *and* $M'[N/x] \twoheadrightarrow (\lambda y.P)^{\alpha}$.

Proof By induction on the length l of $M[N/x] \overset{st}{\twoheadrightarrow} (\lambda y.P)^{\alpha}$.

(i) $l = 0$. Then $M[N/x] = (\lambda y.P)^{\alpha}$. This means that either $M = (\lambda y.M')^{\alpha}$, and $M'[N/x] = P$, or $M = x^{\beta}$ and $x^{\beta}[N/x] = (\lambda y.P)^{\alpha}$.
(ii) $l > 0$. We shall distinguish several subcases, according to the structure of the term M.
   (a) If $M = (\lambda y.M')^{\beta}$, then $M[N/x] = (\lambda y.M'[N/x])^{\beta}$. So, $\alpha = \beta$ and $M'[N/x] \twoheadrightarrow P$.
   (b) If $M = (\ldots((y\ M_1)^{\beta_1}\ M_2)^{\beta_2} \ldots M_n)^{\beta_n}$, then $x = y$, since otherwise $M[N/x]$ could not reduce to a lambda abstraction. So, $M = M'$ and $M'[N/x] \twoheadrightarrow (\lambda y.P)^{\alpha}$.
   (c) If $M = (\ldots((((\lambda y.A)^{\gamma}\ B)^{\beta}\ M_1)^{\beta_1}\ M_2)^{\beta_2} \ldots M_n)^{\beta_n}$, then the redex $((\lambda y.A)^{\gamma}\ B)^{\beta}$ must be contracted along the standard reduction $M[N/x] \overset{st}{\twoheadrightarrow} (\lambda y.P)^{\alpha}$, otherwise we could not get an abstraction as the result of this reduction. Moreover, since $((\lambda y.A)^{\gamma}\ B)^{\beta}$ is the leftmost redex in M, it is the first redex contracted in $M[N/x] \overset{st}{\twoheadrightarrow} (\lambda y.P)^{\alpha}$. Hence, let $Q = (\ldots(((\beta\overline{\gamma} \cdot A[\underline{\gamma} \cdot B/y])\ M_1)^{\beta_1}\ M_2)^{\beta_2} \ldots M_n)^{\beta_n}$. We have $Q[N/x] \overset{st}{\twoheadrightarrow} (\lambda y.P)^{\alpha}$. But the length of this derivation is $l - 1$, so we can apply the induction hypothesis. Since $M \to Q$, we conclude.

□

Lemma 5.3.24 *If $\mathcal{P}$ has an upper bound, and the terms M, N are strongly normalizing, then $M[N/x]$ also is strongly normalizing.*

Proof (We shall abbreviate strongly normalizable to s.n.) Let m be the upper bound for the predicate $\mathcal{P}$; $\mathsf{depth}(M)$ be the length of the longest normalizing derivation of M; and $\|M\|$ be the structural size of M (defined in the obvious way). The proof is by induction on the triple $\ll m - h(\ell(N)), \mathsf{depth}(M), \|M\| \gg$.
(*base case*) By hypothesis, $M = y^{\alpha}$. Then, we have two possibilities: (*i*) $y \neq x$, in which case $M[N/x] = y^{\alpha}$ is trivially s.n.; (*ii*) $x = y$, in which case, $M[N/x] = \alpha \cdot N$ is s.n., for N is s.n.
(*inductive case*) Let us classify subcases according to the size of M:

(i) $M = y^\alpha$. Similar to the base case. In fact, proving it we did not use the hypothesis $m - h(\ell(N)) = 0$.

(ii) $M = (\lambda y.M_1)^\alpha$. Then $M[N/x] = (\lambda y.M_1[N/x])^\alpha$. Obviously, M_1 is s.n., $\mathsf{depth}(M_1) \leq \mathsf{depth}(M)$ and $\|M_1\| < \|M\|$. So, by induction hypothesis, $M_1[N/x]$ is s.n. Thus, $M[N/x]$ is s.n.

(iii) $M = (M_1\ M_2)^\alpha$. Then $M[N/x] = (M_1[N/x]\ M_2[N/x])^\alpha$. Since $\mathsf{depth}(M_1) \leq \mathsf{depth}(M)$ and $\|M_1\| < \|M\|$, $M_1[N/x]$ is s.n. by induction hypothesis. Similarly for $M_2[N/x]$. Now, two subcases are possible:

  (a) $M_1[N/x]$ never reduces to a lambda abstraction. In this case the reduction of $M[N/x]$ is the composition of two independent reductions: a reduction of $M_1[N/x]$ and a reduction of $M_2[N/x]$. So, $M[N/x]$ is s.n. since $M_1[N/x]$ and $M_2[N/x]$ are.

  (b) $M_1[N/x] \twoheadrightarrow (\lambda y.P)^\beta$. The case in which $\mathcal{P}(\beta)$ is false is similar to the previous one. Then, let us assume that $\mathcal{P}(\beta)$ is true. In this case, $M[N/x] = (M_1[N/x]\ M_2[N/x])^\alpha \twoheadrightarrow ((\lambda y.P)^\beta\ M_2[N/x])^\alpha$; furthermore, $\alpha\overline{\beta} \cdot P[\underline{\beta} \cdot M_2[N/x]/y]$ is s.n. Since $M_1[N/x]$ is s.n., we can apply the standardization property of Proposition 5.3.16, getting a standard derivation $M_1[N/x] \overset{\mathsf{st}}{\twoheadrightarrow} (\lambda y.P)^\beta$. Then, by Lemma 5.3.23, we have the following cases:

    1. $M_1 \twoheadrightarrow (\lambda y.M_3)^\beta$ and $M_3[N/x] \twoheadrightarrow P$. Since the term $M = (M_1\ \ M_2)^\alpha$ is s.n., also $M' = \alpha\overline{\beta} \cdot M_3[\underline{\beta} \cdot M_2/y]$ is s.n. Moreover, as $M'[N/x] = \alpha\overline{\beta} \cdot M_3[N/x][\underline{\beta} \cdot M_2[N/x]/y]$, by Lemma 5.3.4 and Lemma 5.3.5, we have

$$M'[N/x] \twoheadrightarrow \alpha\overline{\beta} \cdot P[\underline{\beta} \cdot M_2[N/x]/y]$$

    Moreover, since $M \twoheadrightarrow M'$, $h(\ell(M)) \leq h(\ell(M'))$, and $\mathsf{depth}(M') < \mathsf{depth}(M')$, we can apply induction hypothesis. That is, $M'[N/x]$ is s.n., as well as $\alpha\overline{\beta} \cdot P[\underline{\beta} \cdot M_2[N/x]/y]$.

    2. $M_1 \twoheadrightarrow M_1' = (\dots((x^\gamma A_1)^{\alpha_1}\ A_2)^{\alpha_2} \dots A_n)^{\alpha_n}$ for some M_1' such that $M_1'[N/x] \twoheadrightarrow (\lambda y.P)^\beta$. Since $M_1[N/x]$ is s.n., also P is s.n. Moreover:

$$M_1'[N/x] = (\dots((\gamma \cdot N)\ A_1[N/x])^{\alpha_1} \dots A_n[N/x])^{\alpha_n}$$

    By Lemma 5.3.21, $h(\ell(\gamma \cdot N)) \leq h(\beta)$. Therefore, we have that:

$$h(\ell(N)) \leq h(\ell(\gamma \cdot N)) \leq h(\beta) < h(\underline{\beta}) \leq h(\ell(\underline{\beta} \cdot M_2[N/x]))$$

Then, by induction hypothesis, $\alpha\overline{\beta}\cdot P[\underline{\beta}\cdot M_2[N/x]/y]$ is s.n.

□

Proposition 5.3.25 *If $\mathcal{P}$ has an upper bound, then every term M is strongly normalizable.*

Proof By structural induction on M.

(i) $M = x^\alpha$. Trivial.

(ii) $M = (\lambda y.M_1)^\alpha$. M_1 is strongly normalizing by induction hypothesis, and so is M.

(iii) $M = (M_1\ M_2)^\alpha$. By induction hypothesis, M_1 and M_2 are strongly normalizing. If M_1 never reduces to a lambda abstraction the result is obvious. Otherwise, suppose $M_1 \twoheadrightarrow (\lambda x.M_3)^\beta$; we must prove that $\alpha\overline{\beta} \cdot M_3[\underline{\beta} \cdot M_2/x]$ is strongly normalizing. Since M_3 and M_2 are strongly normalizing, so are $\alpha\overline{\beta} \cdot M_3$ and $\underline{\beta} \cdot M_2$. Thus, by Lemma 5.3.24 we conclude.

□

Theorem 5.3.26 (Church–Rosser) *If $\sigma : M \twoheadrightarrow N$ and $\rho : M \twoheadrightarrow P$, then there exists a term Q such that $N \twoheadrightarrow Q$ and $P \twoheadrightarrow Q$.*

Proof Let $\sigma : M = M_0 \overset{R_1}{\rightarrow} M_1 \overset{R_2}{\rightarrow} \ldots \overset{R_n}{\rightarrow} M_n = N$ and $\rho : M = M_0 \overset{S_1}{\rightarrow} P_1 \overset{S_2}{\rightarrow} \ldots \overset{S_m}{\rightarrow} P_m = P$. Let $\mathcal{P}_\sigma$ and $\mathcal{P}_\rho$ be the predicates defined in the following way:

(i) $\mathcal{P}_\sigma(\alpha)$ if and only if $\exists i, 1 \leq i \leq n, \alpha = \mathsf{degree}(R_i)$;

(ii) $\mathcal{P}_\rho(\alpha)$ if and only if $\exists j, 1 \leq j \leq m, \alpha = \mathsf{degree}(S_j)$.

Let $\mathcal{P} = \mathcal{P}_\sigma \bigcup \mathcal{P}_\rho$. Obviously, $\mathcal{P}$ has an upper bound, and since the two reductions σ and ρ are legal for $\mathcal{P}$, M is strongly normalizable (with respect to P). Then, by Proposition 5.3.10, we conclude. □

Theorem 5.3.27 (standardization) *If $M \twoheadrightarrow N$, then $M \overset{st}{\twoheadrightarrow} N$.*

Proof Similar to the previous one, using Proposition 5.3.16 in the place of Proposition 5.3.10. □

Exercise 5.3.28 (difficult) Prove that if $\sigma : M \twoheadrightarrow N$ and $\rho : M \twoheadrightarrow P$, then there exists a term Q such that $\rho/\sigma : N \twoheadrightarrow Q$ and $\sigma/\rho : P \twoheadrightarrow Q$.

Exercise 5.3.29 (difficult) Prove that if $\rho : M \twoheadrightarrow N$, then there exists a standard reduction $\rho_s : N \overset{st}{\twoheadrightarrow} Q$ such that $\rho \equiv \rho_s$.

Remark 5.3.30 In the proof of the Church–Rosser property (Theorem 5.3.26) we have not explicitly seen how closing reductions are done. Anyhow, the previous exercise shows that, similarly to when proving local confluence, the diamond is closed by the image of the reduction on the opposite side. Furthermore, let us note that this property is in some sense encoded into the proof of Theorem 5.3.26: the predicate $\mathcal{P}$ that we choose is exactly the one that enables families that were present in the initial reductions ρ and σ.

5.3.2 Labeled and unlabeled λ-calculus

We shall now prove one of the most interesting results of the labeled λ-calculus. As we noted when studying families, in the unlabeled case there are pairs of reductions starting and ending with the same pair of terms that cannot be considered equivalent (recall the example $(I\,(I\,x))$). We pointed out with an example that this is not the case in the labeled calculus. In this section we will prove that this is a general property of labeled λ-calculus. In fact, we will see that given two terms M and N such that $M \twoheadrightarrow N$, then there is a unique standard reduction $M \overset{st}{\twoheadrightarrow} N$.

Definition 5.3.31 Given a labeled term M, we call $\tau^{-1}(M)$ the corresponding unlabeled term obtained by erasing all labels. Formally:

(i) $\tau^{-1}(x^\alpha) = x$;
(ii) $\tau^{-1}((\lambda x.M)^\alpha) = \lambda x.\tau^{-1}(M)$;
(iii) $\tau^{-1}((M\ N)^\alpha) = (\tau^{-1}(M)\ \tau^{-1}(N))$.

We shall now introduce a new measure for labels that will become useful in proving the next proposition.

Definition 5.3.32 The *size* $\|\alpha\|$ of a label α is the sum of the total number of letters, overlinings and underlinings in α. Formally:

(i) $\|a\| = 1$ if a is an atomic label;
(ii) $\|\alpha\beta\| = \|\alpha\| + \|\beta\|$;
(iii) $\|\overline{\alpha}\| = \|\underline{\alpha}\| = \|\alpha\| + 1$.

Lemma 5.3.33 *If* $M \twoheadrightarrow M'$, *then* $\|\ell(M)\| \leq \|\ell(M')\|$.

Proof Immediate. □

Proposition 5.3.34 *For every pair of labeled terms* U, V *such that* $U \twoheadrightarrow V$, *there exists exactly one standard reduction* $U \overset{st}{\twoheadrightarrow} V$.

Proof If $U \twoheadrightarrow V$, by the standardization theorem there exists at least one standard reduction $\sigma : U \twoheadrightarrow V$. We must prove that this is unique. The proof is by induction on the length l of σ and the structure of U.
(*base case*) $U = V = x^\alpha$. Then the only standard reduction is the empty one.
(*inductive case*)

(i) $U = x^\alpha$. This case has been already considered.

(ii) $U = (\lambda x.U_1)^\alpha$. Then $V = (\lambda x.V_1)^\alpha$ and $U_1 \overset{st}{\twoheadrightarrow} V_1$ with a standard reduction σ' of length l. By induction hypothesis σ' is unique, and so is σ.

(iii) $U = (U_1\ U_2)^\alpha$. σ must be either of the kind

$$U = (U_1\ U_2)^\alpha \overset{st}{\twoheadrightarrow} (V_1\ U_2) \overset{st}{\twoheadrightarrow} (V_1\ V_2)^\alpha = V$$

or

$$U = (U_1\ U_2)^\alpha \overset{norm}{\twoheadrightarrow} ((\lambda x.U_3)^\beta\ U_2) \to \alpha\overline{\beta} \cdot U_3[\underline{\beta} \cdot U_2/x] \overset{st}{\twoheadrightarrow} V$$

where $(\lambda x.U_3)^\beta = \mathsf{Abs}(U_1)$.
We prove first that, in the labeled system, *all* standard reductions between U and V are of a same kind. Indeed, suppose we have two standard reductions ρ and τ of different kinds. If ρ is of the first kind above, we also have $\ell(V) = \alpha = \ell(U)$. On the other side, if τ is of the second kind, by Lemma 5.3.33, we have

$$\|\ell(U)\| = \|\alpha\| < \|\ell(\alpha\overline{\beta} \cdot U_3[\underline{\beta} \cdot U_2/x])\| \leq \|\ell(V)\|$$

and thus $\ell(U) \neq \ell(V)$.
Coming back to the proof of the proposition, we can thus distinguish two subcases, according to the kind of standard reduction σ.

(a) If σ is of the first kind, the result follows by induction, since the standard reductions $U_1 \overset{st}{\twoheadrightarrow} V_1$ and $U_2 \overset{st}{\twoheadrightarrow} V_2$ are unique.

(b) If σ is of the second kind, every other standard reduction must be of the same kind. Since the initial part of the derivation is normal, this must be common to every such reduction. Since, by induction on the length of the derivation, there exists a unique standard reduction from $\alpha\overline{\beta} \cdot U_3[\underline{\beta} \cdot U_2/x]$ to V, σ is unique too.

$\square$

Note that the previous proposition does not hold in the (unlabeled) λ-calculus (take for instance $U = (I(I\ x))$ and $V = (I\ x)$). We also invite the reader no to confuse the previous property with the one that Theorem 5.3.27 induces on the unlabeled λ-calculus. In fact, in the unlabeled calculus we can still say that given a reduction $\rho : M \twoheadrightarrow N$, there is a unique standard reduction $\rho_s : M \overset{st}{\twoheadrightarrow} N$ equivalent to it (see also Exercise 5.3.29). Nevertheless, two non-equivalent reductions connecting the same pair of terms lead to two distinct standard reductions. This is no longer true in the labeled calculus where two reductions are equivalent if and only if they connect the same pair of terms (see next theorem). In a sense, the labeled λ-calculus *does not make syntactical mistakes* identifying terms coming from 'different' derivations. This intuition is formalized by the following theorem.

Theorem 5.3.35 *Let* U *be a labeled* λ*-term with* $M = \tau^{-1}(U)$*. Let* $\sigma : M \twoheadrightarrow N$ *and* $\rho : M \twoheadrightarrow P$*. Consider the labeled reductions* $\sigma_1 : U \twoheadrightarrow V$ *and* $\rho_1 : U \twoheadrightarrow W$ *respectively isomorphic to* σ *and* ρ*. Then:*

(i) $\sigma \equiv \rho$ *if and only if* $V = W$*;*

(ii) $\sigma \sqsubseteq \rho$ *if and only if* $V \twoheadrightarrow W$*.*

Proof

(i) Let σ_1' and ρ_1' be the standard reductions corresponding to σ_1 and ρ_1, respectively. Then, $\sigma \equiv \tau$ if and only if $\sigma_1 \equiv \tau_1$ if and only if $\sigma_1' = \tau_1'$. But for the previous proposition, $\sigma_1' = \tau_1'$ if and only if $V = W$.

(ii) By definition, $\sigma \sqsubseteq \rho$ if and only if there exists τ such that $\sigma\tau \equiv \rho$. Let $\tau_1 : V \twoheadrightarrow T$ be the labeled reduction isomorphic to τ. By the previous item, $\sigma\tau \equiv \rho$ if and only if $T = W$. So, $\sigma \sqsubseteq \tau$ if and only if $V \twoheadrightarrow W$.

$\square$

5.3.3 Labeling and families

Proposition 5.3.36 *If* $M \twoheadrightarrow N$ *in the labeled lambda calculus, and* S *is a redex in* M*, then all residuals of* S *in* N *have the same degree as* S.

Proof Let us start with the case $M \xrightarrow{R} N$. Let $R = ((\lambda x.A)^\alpha\ B)^\beta$ and $S = ((\lambda y.C)^\gamma\ D)^\delta$. The proof is by cases, according to the mutual positions of R and S in M.

(i) R and S are disjoint. Trivial.

(ii) If S contains R in C (respectively D), then the (unique) residual of S in N is of the form $((\lambda y.C')^\gamma\ D)^\delta$ (respectively $((\lambda y.C)^\gamma\ D')^\delta$, which has the same degree as S.

(iii) If R contains S in A, then S has a unique residual in N of the form $((\lambda y.C[\underline{\alpha} \cdot B/x])^\gamma\ D[\underline{\alpha} \cdot B/x])^{\delta'}$, where only the external label δ can be modified into δ' (this happens when $A = ((\lambda y.C)^\gamma\ D)^\delta$).

(iv) If R contains S in B, then all residuals of the redex S in N are of the form $((\lambda y.C)^\gamma\ D)^{\delta'}$, where only the external label δ can be modified into δ' (this happens when $B = ((\lambda y.C)^\gamma\ D)^\delta$).

□

Exercise 5.3.37 Reformulate the proof of the previous proposition in terms of labeled syntax trees. Note in particular that the four cases in the proof correspond in order to: (*i*) the edge of the redexes appear in disjoint subtrees; (*ii*) the edge of R is in the subtree of the application corresponding to S; (*iii*) the edge of S is in the body of the abstraction of R; (*iv*) the edge of S is in the subtree of the argument of R. In particular, note that the last one is the only case in which the edge of R is duplicated.

Proposition 5.3.38 *If* $M \xrightarrow{R} N$ *and* S *is a redex in* N *created by* R *then* $\|\mathsf{degree}(R)\| < \|\mathsf{degree}(S)\|$.

Proof Let $R = ((\lambda x.A)^\alpha\ B)^\beta$. Three cases are possible:

Upward creation. In M there is a subterm $(((\lambda x.A)^\alpha\ B)^\beta\ D)^\delta$ where $A = (\lambda y.C)^\gamma$. Then, $S = ((\beta\overline{\alpha} \cdot A[\underline{\alpha} \cdot B/x])\ D)^\delta = ((\lambda y.C[\underline{\alpha} \cdot B/x])^{\beta\overline{\alpha}\gamma}\ D)^\delta$ and $\|\alpha\| < \|\beta\overline{\alpha}\gamma\|$.

Downward creation. In A there is a subterm of the form $((x)^\epsilon\ D)^\delta$ and $B = (\lambda y.C)^\gamma$. Then, $S = ((x)^\epsilon\ D)^\delta[\underline{\alpha} \cdot (\lambda y.C)^\gamma/x] = ((\lambda y.C)^{\epsilon\underline{\alpha}\gamma}\ D)^\delta$, and $\|\alpha\| < \|\epsilon\underline{\alpha}\gamma\|$.

Identity. This is a combination of the two previous cases. Namely, we have a subterm in M of the form $(((\lambda x.A)^{\alpha}\ B)^{\beta}\ D)^{\delta}$ where $A = x^{\epsilon}$ and $B = (\lambda y.C)^{\gamma}$. Then, $S = ((\beta\overline{\alpha} \cdot A[\underline{\alpha} \cdot B/x])\ D)^{\delta} = ((\lambda y.C)^{\beta\overline{\alpha}\epsilon\underline{\alpha}\gamma}\ D)^{\delta}$ and $\|\alpha\| < \|\beta\overline{\alpha}\epsilon\underline{\alpha}\gamma\|$.

□

Definition 5.3.39 (INIT) We shall say that the predicate $\mathsf{INIT}(M)$ is verified by a labeled term M if and only if the labels of all subterms of M are atomic and pairwise distinct.

Proposition 5.3.40 *Let M be a term such that $\mathsf{INIT}(M)$ holds. For any reduction $M \twoheadrightarrow N$, a redex S in N is a residual of a redex R in M if and only if they have the same degree.*

Proof The only if direction is Proposition 5.3.36. For the if direction, we proceed by induction on the length l of the reduction $M \twoheadrightarrow N$. If $l = 0$, the result follows by the conditions imposed by $\mathsf{INIT}(M)$. If $l > 0$, let $M \twoheadrightarrow N' \xrightarrow{P} N$. By hypothesis, there exists a redex R in M with the same degree as S. Since, by $\mathsf{INIT}(M)$, all redexes in M have atomic degree, the degree of S must be atomic too. This means that S cannot be created by the firing of P, since otherwise $\|\mathsf{degree}(S) > 1\|$. So, S must be the residual of a redex S' in M' with the same degree as R. By induction hypothesis S' is a residual of R, and so is S. □

Theorem 5.3.41 *Let ρR and σS be two redexes with histories $\rho : M \twoheadrightarrow N$ and $\sigma : M \twoheadrightarrow P$. Let us take the corresponding isomorphic reductions $\rho_1 : U \twoheadrightarrow V$ and $\sigma_1 : U \twoheadrightarrow W$ in the labeled system (the initial labeling of U can be arbitrary). If $\rho R \simeq \sigma S$, then R and S have the same degree (in V and W, respectively).*

Proof This is an easy corollary of Theorem 5.3.35 and Proposition 5.3.36.

□

Let us remark again that the previous theorem holds for any labeling of the initial term M. According to the intended interpretation of labeling as the name associated to all the edges with the same origin, the previous theorem states that this interpretation is sound. In fact, all redexes in the same family with respect to a given initial term are marked by the same label.

In Chapter 6 we shall prove the converse of Theorem 5.3.41, but under the (essential) assumption that $\mathsf{INIT}(U)$ holds. In fact, according to the idea that all redexes in the initial term are in different families and then unsharable, all the edges of the initial term must be marked with different (atomic) labels—in other words, the initial term is not the result of some previous computation or equivalently, we do not know its history. Under this assumption, we will be able to prove that two redexes are in the same family only if they have the same degree. Let us note that Proposition 5.3.40 is not enough to give this equivalence between the equivalence families and degree. In fact, Proposition 5.3.40 proves the result in the case of atomic labels, but does not allow us to immediately extend it to the case of composite degrees. Let us compare this with the situation in extraction relation. Also in that case it is immediate to prove that when a redex R has an ancestor S in the initial term, then S is the canonical representative of R plus its history (see Lemma 5.2.8). The difficult part in proving uniqueness of a canonical representative is when the redex is created along the reduction (see the rest of section 5.2). In the labeled case, the difficult part will be to show the uniqueness of this canonical representative in terms of labels, *i.e.*, that two canonical representatives for extraction cannot have the same degree.

5.4 Reduction by Families

In this section we aim to find the syntactic counterpart of an evaluator never duplicating redexes, according to the notion of duplication induced by the family relation (*i.e.*, the zig-zag relation defined in Section 5.1). We introduce a strategy of derivation by families, a parallel reduction in which at each step several redexes in the same family can be reduced in parallel. The idea is that by systematically reducing in parallel *all* the redexes in a given family, another member of that family cannot appear later on during the computation. This follows from the interpolation property of Lemma 5.1.12.

In order to formalize these ‘parallel’ derivations, we first generalize the finite development theorem.

Definition 5.4.1 Let $[\rho R] = \{\sigma S \mid \sigma S \simeq \rho R\}$ be the *family class* of ρR. Let $\rho = \mathcal{F}_1 \cdots \mathcal{F}_n$. Then,

$$\mathsf{FAM}(\rho) = \{[\mathcal{F}_1 \cdots \mathcal{F}_i R] \mid R \in \mathcal{F}_{i+1}, i = 0, 1, \ldots, n-1\}$$

is the set of family classes contained in ρ.

Definition 5.4.2 (development of family classes) Let $\mathcal{X}$ be a set of family classes. A derivation ρ is *relative* to $\mathcal{X}$ if $\mathsf{FAM}(\rho) \subseteq \mathsf{X}$. A derivation ρ is a *development* of $\mathcal{X}$ if there is no redex R such that $[\rho\mathsf{R}] \in \mathcal{X}$.

Theorem 5.4.3 (generalized finite developments) *Let $\mathcal{X}$ be a (finite) set of family classes. Then:*

(i) *There is no infinite derivation relative to $\mathcal{X}$.*

(ii) *If ρ and σ are two developments of $\mathcal{X}$ then, $\rho \equiv \sigma$.*

Proof

(i) Let $\mathcal{X}$ be a finite set of family classes with respect to the initial term M. Let us assume that $\mathsf{INIT}(\mathsf{M})$ is true. Let $\Gamma = \{\mathsf{degree}(\rho\mathsf{R}) \mid [\rho\mathsf{R}] \in \mathcal{X}\}$ and $\mathcal{P} = \{\alpha \mid \mathsf{h}(\alpha) \leq \mathsf{max}\{\mathsf{h}(\beta) \mid \beta \in \Gamma\}\}$. By Theorem 5.3.41, any family class has a unique degree. Hence, Γ is finite and $\mathcal{P}$ has an upper bound. By Proposition 5.3.25, any labeled derivation legal for $\mathcal{P}$ is finite. Since any derivation relative to $\mathcal{X}$ is legal for $\mathcal{P}$, we conclude.

(ii) Let us observe that, if ρ and σ are relative to $\mathcal{X}$, then $\rho \sqcup \sigma$ is also relative to $\mathcal{X}$. Now, by definition, if ρ and σ are two developments of $\mathcal{X}$, then $\rho \equiv \rho \sqcup \sigma$ and $\sigma \equiv \rho \sqcup \sigma$. Thus, $\rho \equiv \sigma$ by transitivity.

□

Let us now prove our claim that, after the development of a set of families $\mathcal{X}$, no redex in a family contained in $\mathcal{X}$ can be created along the reduction. The main lemma is preceded by a more technical property useful for its proof.

Lemma 5.4.4 *Let $\rho\mathsf{R}$ be the canonical derivation of $\sigma\mathsf{S}$. Then $\rho \sqsubseteq \sigma$ if and only if $\rho\mathsf{R} \leq \sigma\mathsf{S}$.*

Proof The if direction follows by definition of copy relation. Let us focus on the only-if direction. Let τ be the standard derivation of σ. Then $\tau\mathsf{S} \rhd \rho\mathsf{R}$ and $\rho \sqsubseteq \tau$, since $\tau \equiv \sigma$. Furthermore, it suffices to prove $\rho\mathsf{R} \leq \tau\mathsf{S}$. We proceed by induction on $|\tau|$.

If $\tau = 0$, then $\mathsf{S} \rhd \rho\mathsf{R}$ implies $\rho = 0$ and $\mathsf{R} = \mathsf{S}$. Thus, $\rho\mathsf{R} \leq \tau\mathsf{S}$.

Let $\tau = \mathsf{T}\tau'$. Since both ρ and τ are standard, there are two cases:

(i) $\rho = T\rho'$ and $\rho' R$ is the canonical derivation of $\tau' S$. Then $\rho \sqsubseteq \tau$ implies $\rho' \sqsubseteq \tau'$ by left cancellation. Therefore, $\rho R \leq \tau S$ because, by induction, $\rho' R \leq \tau' S$.

(ii) $T\rho' R' \rhd \rho R$, where $\rho' R'$ is the canonical derivation of $\tau' S$. Then $\rho' \sqsubseteq \rho/T$, by definition of $\rhd$. This fact and $\rho \sqsubseteq \tau$ imply $\rho/T \sqsubseteq \tau/T = \tau'$. Therefore, $\rho' \sqsubseteq \rho/T \sqsubseteq \tau'$ and, by induction, $\rho' R' \leq \tau' S$. By the interpolation property (Lemma 5.1.12), there is S' such that $\rho' R' \leq (\rho/T)S' \leq \tau' S$. Then, $T(\rho/T)S' \leq (T\tau')S = \tau S$. Moreover, $\rho(T/\rho)S' \leq \tau S$, for $T(\rho/T) \equiv \rho(T/\rho)$. Finally, $S \in R/(T/\rho)$, since ρR is canonical and $\rho(T/\rho)S' \simeq \rho R$. Thus, $\rho R \leq \tau S$.

□

Lemma 5.4.5 *Let* ρ *be a development of* $\mathcal{X}$*.*

(i) *Let* σS *be the canonical derivation of* ρR*. Then* $\sigma S \leq \rho R$*.*

(ii) *For every* σS *such that* $\rho \sqsubseteq \sigma$*,* $[\sigma S] \notin \mathsf{FAM}(\rho)$*.*

Proof

(i) By hypothesis, σ is relative to $\mathcal{X}$. Then $\sigma \sqsubseteq \rho$, since σ can always be extended to a development $\sigma\tau$ of $\mathcal{X}$ and $\sigma\tau \equiv \rho$, by Theorem 5.4.3. Thus $\sigma S \leq \rho R$, by Lemma 5.4.4.

(ii) By contradiction. Let $\sigma S \in \mathsf{FAM}(\rho)$, $\rho = \rho_1 \mathcal{F} \rho_2$ and $R \in \mathcal{F}$ such that $\rho_1 R \simeq \sigma S$. Let $\rho' R'$ be the canonical derivation of $\rho_1 R$ and σS. By definition, ρ' is relative to $\mathsf{FAM}(\rho)$. Therefore, $\rho' \sqsubseteq \rho$ and, by transitivity, $\rho' \sqsubseteq \sigma$. Therefore, $\rho' R' \leq \sigma S$, by Lemma 5.4.4. By the interpolation property (Lemma 5.1.12), there exists T such that $\rho' R' \leq \rho T \leq \sigma S$. But this contradicts the hypothesis that ρ is a development of $\mathcal{X}$, since under this hypothesis ρ should also be a development of $\mathsf{FAM}(\rho)$.

□

5.4.1 Complete derivations

We can finally define the reductions by families we are interested in.

Definition 5.4.6 (complete derivation) We say that the derivation $\mathcal{F}_1 \cdots \mathcal{F}_n$ is *complete* if and only if $\mathcal{F}_i \neq \emptyset$ and $\mathcal{F}_i$ is a maximal set of redexes such that

$$\forall R, S \in \mathcal{F}_i.\ \mathcal{F}_1 \cdots \mathcal{F}_{i-1} R \simeq \mathcal{F}_1 \cdots \mathcal{F}_{i-1} S$$

for $i = 1, 2, \dots, n$.

The following lemma shows that complete derivations are particular developments.

Lemma 5.4.7 *Every complete derivation* ρ *is a development of* $\mathsf{FAM}(\rho)$.

Proof By induction on $|\rho|$. The base case is obvious. Let $\rho = \sigma\mathcal{F}$, where $\mathcal{F}$ is a maximal set of redexes in the same family. Then, by inductive hypothesis, σ is a development of $\mathsf{FAM}(\sigma)$. Therefore, by Lemma 5.4.5(ii), there is no $\sigma'S$, $\sigma \sqsubseteq \sigma'$, such that $[\sigma'S] \in \mathsf{FAM}(\sigma)$. By contradiction, let us assume that there exists R such that $\rho R \simeq \sigma S$. Namely, let us assume that ρ is not a development of $\mathsf{FAM}(\rho)$. Let $\tau S'$ be the canonical derivation of σS. Then $\tau \sqsubseteq \sigma \sqsubseteq \rho = \sigma\mathcal{F}$. Moreover, by Lemma 5.4.5(i), $\tau S' \leq \rho R$ and, by the interpolation lemma (Lemma 5.1.12), there exists T such that $\sigma T \leq \rho R$. This means that $\sigma T \simeq \sigma S$ with $R \in T/\mathcal{F}$, invalidating the hypothesis that $\mathcal{F}$ is a maximal set of redexes in the same family. □

Complete derivations contract *maximal* set of copies of a single redex. When complete derivations are considered, this means that deciding if two redexes are in the same family may be safely reduced to checking the copy-relation.

Lemma 5.4.8 *A derivation* $\rho = \mathcal{F}_1 \cdots \mathcal{F}_n$ *is complete if and only if, for* $i = 1, \dots, n$, $\mathcal{F}_i$ *is a* maximal set of copies. *Namely, for any* i, *there exist* $\sigma_i S_i$ *and* τ_i *such that* $\sigma_i \tau_i \equiv \rho_i$ *and* $\mathcal{F}_i = S/\tau_i$.

Proof

(*only-if direction*) Let ρ be a complete reduction and R be a redex with history ρ. Let $\mathcal{F}$ be the set of redexes T such that $\rho R \simeq \rho T$. Let $\mathcal{F}'$ be a set of redexes containing R and such that there exist σS and τ with $\sigma\tau \equiv \rho$ and $\mathcal{F}' = S/\tau$. We shall prove that $\mathcal{F} = \mathcal{F}'$. Surely, $\mathcal{F}' \subseteq \mathcal{F}$, by definition of $\simeq$. By the completeness of ρ and Lemma 5.4.7, ρ is a development of $\mathsf{FAM}(\rho)$. Let $\rho'R'$ be the canonical derivation of ρR. Then $\rho'R' \leq \rho R$, by Lemma 5.4.5(i). Therefore, for every $S \in \mathcal{F}$, $\rho'R' \leq \rho S$; which means $\mathcal{F} \subseteq \mathcal{F}'$, for $\mathcal{F}'$ is maximal.

(*if direction*) By contradiction. Let us assume that ρ is complete and that there exists i such that $\mathcal{F}_i$ is not a maximal set of copies. Let $R \in \mathcal{F}_i$. Therefore, there is $S \notin \mathcal{F}_i$ such that, for some τT, $\tau T \leq \mathcal{F}_1 \cdots \mathcal{F}_{i-1} R$ and $\tau T \leq \mathcal{F}_1 \cdots \mathcal{F}_{i-1} S$. Then, $\mathcal{F}_1 \cdots \mathcal{F}_{i-1} R \simeq \mathcal{F}_1 \cdots \mathcal{F}_{i-1} S$, which invalidates the hypothesis that $\mathcal{F}_i$ is a maximal set of redexes in the same family.

□

Exercise 5.4.9 Prove that for any reduction $\rho : M \twoheadrightarrow N$ there exists a complete reduction ρ_c equivalent to ρ (*i.e.*, $\rho_c \equiv \rho$).

Proposition 5.4.10 *Let* ρ *be a complete derivation. We have that* $|\rho| = \sharp(\mathsf{FAM}(\rho))$, *where* $\sharp(\mathsf{FAM}(\rho))$ *is the cardinality of* $\mathsf{FAM}(\rho)$.

Proof Easy consequence of Lemma 5.4.7 and of the requirement that the steps of complete derivations are non-empty. □

Reasoning in graphs, it will be useful to have names for each link participating in a β-reduction.

Definition 5.4.11 In β-reduction, the links that are consumed by and created by reducing a redex will be referred to as indicated in Figure 5.8.

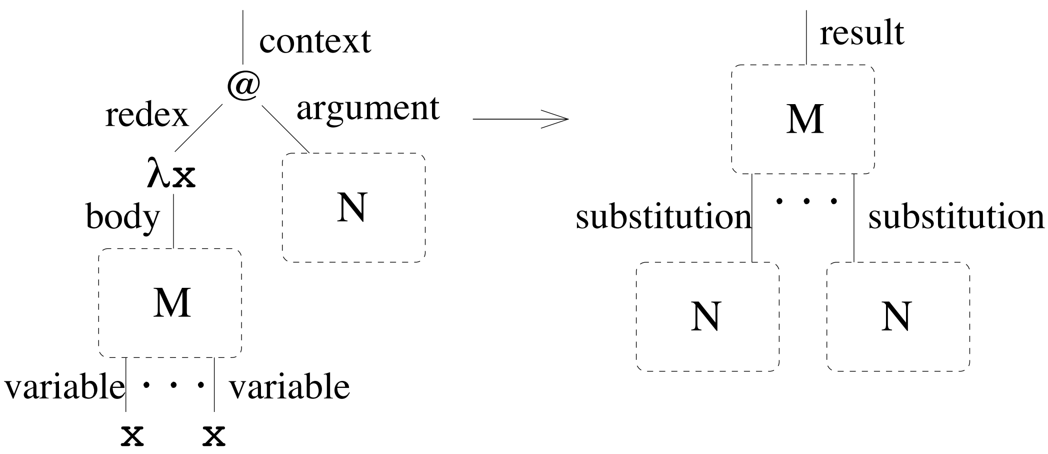


Fig. 5.8. Links involved in β-reduction.

Lemma 5.4.12 *Let* $\mathcal{F}_1 \cdots \mathcal{F}_n$ *be a complete derivation. Then:*

(i) *Each new link created by a (parallel) reduction is marked with a label that did not previously appear in the expression.*

(ii) *The labels on result links created by* $\mathcal{F}_i$ *are different from the labels on substitution links. The labels on two result links created by* $\mathcal{F}_i$ *are identical if and only if the labels on the antecedent context, redex and body links are respectively equal. The labels* *on two substitution links created by $\mathcal{F}_i$ are identical if and only if the labels on the antecedent variable, redex and argument links are respectively equal.*

Proof Easy, by definition of labeling and complete derivation. □

5.5 Completeness of Lamping's Algorithm

Lévy's work concluded in 1980 leaving open the issue of designing a λ-evaluator implementing complete reductions. This issue lay dormant for ten years before Lamping awakened it presenting its evaluator in 1989.

Having the formalization of Lévy optimality at hand we can now prove that the algorithm presented in Chapter 3—a simplified version of Lamping's theorem—fits Lévy's completeness requirement. Most of the definitions and results in this section are due to Lamping [Lam89].

In order to prove that Lamping's algorithm is complete, it suffices to show that all the redexes in a maximal set have a unique representation in the sharing graphs. Let us be more precise on this point.

Let G be a sharing graph obtained along the reduction of $[M]$. Let T be the λ-term that matches with G (see Definition 3.5.13). Correctness implies that $\rho : M \twoheadrightarrow T$ (see Theorem 3.5.15). Furthermore, let ρ_c be a complete reduction equivalent to ρ (see Exercise 5.4.9), we have to prove that in G all the redexes in the same family have a unique representation, in this way contracting such a shared redex we would contract a maximal set of redexes. In terms of labels, to prove completeness means showing that any β-redex edge in G represents *only and all* β-redexes of T with the same label.

The proof will be pursued exploiting the latter correspondence of labels. Nevertheless, it is not trivial, for the correspondence between edges and connections in the sharing graph is not a function from labels to identity connections (*i.e.*, sequences of edges crossing fans or brackets only). Namely, uniqueness of representation might not hold for edges with the same label that are not β-redexes.

For instance, let us consider the example in Figure 5.9. The edge marked u in the sharing graph in Figure 5.9(1) represents a set of edges in the syntax tree associated to the graph, all with the same label. However, as soon as the fan-@ interaction is contracted (see Figure 5.9(2)), the edge u is split into the two (distinct) paths $u' \cdot v$ and $u' \cdot w$. The representations of these paths in the syntax tree are the same as u, since a fan-@ interaction does not change the syntax tree matching the sharing