Normalisation by Random Descent*

Vincent van Oostrom[1] and Yoshihito Toyama[2]

1 **Department of Computer Science, University of Innsbruck, Innsbruck, Austria**
`Vincent.van-Oostrom@uibk.ac.at`
2 **RIEC, Tohoku University, Sendai, Japan**
`toyama@riec.tohoku.ac.jp`

Abstract

We present abstract hyper-normalisation results for strategies. These results are then applied to term rewriting systems, both first and higher-order. For example, we show hyper-normalisation of the left–outer strategy for, what we call, left–outer pattern rewrite systems, a class comprising both Combinatory Logic and the $\lambda\beta$-calculus but also systems with critical pairs. Our results apply to strategies that need not be deterministic but do have Newman's random descent property: all reductions to normal form have the same length, with Huet and Lévy's external strategy being an example. Technically, we base our development on supplementing the usual notion of commutation diagram with a notion of order, expressing that the measure of its right leg does not exceed that of its left leg, where measure is an abstraction of the usual notion of length. We give an exact characterisation of such global commutation diagrams, for pairs of reductions, by means of local ones, for pairs of steps, we dub Dyck diagrams.

1998 ACM Subject Classification F.3.1 Specifying and Verifying and Reasoning about Programs, F.4.1 Mathematical Logic, F.4.2 Grammars and Other Rewriting Systems

Keywords and phrases strategy, hyper-normalisation, commutation, random descent

Digital Object Identifier 10.4230/LIPIcs.FSCD.2016.32

1 Introduction

The (hyper-)normalisation of the leftmost–outermost strategy is a fundamental result in Combinatory Logic and the λ-calculus, cf. [3, 7]. For the special case of the λ-calculus, the simple idea underlying the present paper is that normalisation of the leftmost–outermost strategy is due to it being both *deterministic*, there is at most one leftmost–outermost step from any given λ-term, and *compatible* with β, in the sense that if $M \to_\beta N$ then repeatedly performing the leftmost–outermost strategy on both M and N results either in a common reduct or in infinite reductions from both. Compatibility guarantees (Section 5) that each term β-convertible to some normal form is also convertible to that normal form by leftmost–outermost steps. Determinism guarantees that if there is such a *conversion* to normal form, then there *exists* a leftmost–outermost *reduction* from the term to the normal form, and so *all* leftmost–outermost reductions from that term terminate. A method for proving hyper-normalisation is obtained from this by strengthening compatibility with an *order* constraint expressing that in the above the leftmost–outermost reduction from M be at least as long as that from N. Ordered compatibility guarantees that β-steps never increase the *distance*, i.e. the length of the leftmost–outermost reduction of a term to its normal

* The research for this paper was initiated during a visit in November 2013 funded by RIEC, Tohoku University, of the first author to the second author.

form, from which hyper-normalisation follows as leftmost–outermost steps do decrease that distance. We present an abstract account of this idea based on the following observations.

The first observation (Section 4) is that determinism can be relaxed to Newman's *random descent* property, all reductions to normal form have the same length [13, 20, 21, 16]. This vastly broadens the scope of the method. In addition to deterministic strategies such as leftmost–outermost, it covers, e.g., interaction net strategies, linear β-reduction, and Huet and Lévy's external strategy [8], allowing to contract redexes that are outermost until eliminated (by contracting the redex itself or some overlapping redex [19]). Technically, whereas for deterministic systems the notion of distance is well-defined because given an object there is a *unique reduction* to its normal form, it is still well-defined for random descent systems since although reductions to normal form then need no longer be unique they have a *unique length*.

The second observation (Section 3) is that all reductions from an object to normal form having the same length is equivalent to the *order* constraint that for each pair of reductions from the object to normal form, the first is at least as long as the second. Working towards deciding it, we characterise this property of peaks of reductions by means of a property of *local* peaks (of steps). More precisely, we show it necessary and sufficient that such local peaks be completable by means of a *Dyck*-conversion, a conversion in which the number of forward steps never (for any prefix) exceeds the number of backward steps. Technically, we establish the above in a *commutation* setting where the reductions being ordered may be reductions in two distinct rewrite systems. In Section 6 we turn this local criterion into a *critical pair* criterion for a concrete class of higher-order term rewrite systems, dubbed *left–outer Dyck*, comprising both Combinatory Logic and the $\lambda\beta$-calculus.

The third observation is that we may abstract the notion of length in the random descent property into a notion of measure, allowing steps having different measures to coexist. This covers, for example, systems having 'macro' steps abbreviating several 'micro' steps. We build this into our setup right from the start (Section 2), by introducing *derivation* and *conversion* monoids which allow to measure reductions respectively conversions.

For terminating systems, the classical division of work for proving confluence is to first *localise* the confluence property for abstract rewrite systems (Newman's Lemma), and then using that establish confluence for term rewrite systems by means of a *critical pair* criterion (Huet's Critical Pair Lemma). We structure our paper accordingly, first localising random descent to the *local Dyck property* (Sections 3, 4) for abstract rewrite systems, and then using that to establish a critical pair criterion, the *left–outer Dyck property*, for establishing hyper-normalisation of the left–outer strategy for term rewrite systems (Sections 5, 6).

We employ the untyped λ-calculus with β- and/or η-reduction [2] as a running example, marking the examples where en passant new results are obtained by a double dagger (‡).

Contribution. Apart from unifying our earlier results [20, 21, 16], with a clean separation into results for abstract and term rewriting, the main contributions of this paper are the notions of measured rewrite system, ordered commutation, compatibility and Dyck diagrams.

2 Preliminaries

We assume basic knowledge of term rewriting and the λ-calculus [1, 2, 19], and use notation from [19]. We employ abstract rewrite systems (ARSs) and strategies for them as defined in [19, Chapters 8 and 9]; cf. also [16]. In particular, *steps* are first-class citizens of abstract rewrite systems[1] and a *strategy* for a rewrite system is a sub-rewrite system on the same set of objects, with the same set of normal forms.[2] Throughout we assume

$\twoheadrightarrow$, $\to$, ${\color{red}\twoheadrightarrow}$, ${\color{blue}\rightarrowtriangle}$ are rewrite systems on the same set of objects, unless stated otherwise.

A strategy $\to$ for the rewrite system $\twoheadrightarrow$ is ($\twoheadrightarrow$-)*normalising* if $\to$ is terminating on objects that are $\twoheadrightarrow$-normalising, i.e. if for all a with $a \twoheadrightarrow^* b$ for some $\twoheadrightarrow$-normal form b, a is $\to$-terminating. The strategy $\to$ is $\twoheadrightarrow$*-hyper-normalising* if $\to/\twoheadrightarrow$, i.e. $\twoheadrightarrow^* \cdot \to \cdot \twoheadrightarrow^*$, is $\twoheadrightarrow$-normalising.

▶ **Example 1.** For the $\lambda\beta$-calculus, selecting an outermost redex for contraction yields a strategy (since any term not in β-normal form contains some outermost redex) that is non-deterministic, the *outermost* strategy. Starting from the outermost strategy, selecting the *leftmost* redex (in the tree ordering, cf. Definition 27) for rewriting yields a strategy for it that is deterministic. Composing both yields a strategy for β-reduction that is deterministic again, the *leftmost–outermost* strategy. The leftmost strategy starting from β-reduction, is also well-defined but non-deterministic. The leftmost–outermost strategy is hyper-normalising [3], but the leftmost and outermost strategies are not even normalising, cf. $(\lambda x.x\Omega)(\lambda xy.z)\Omega$.

For convenience we recapitulate notational conventions and their mnemonic values.

▶ **Notation 2.** *We employ:*

- $\twoheadrightarrow$ *arrow notation to denote an ARS having steps as first-class citizens;*
- $\to$ *part of notation* $\twoheadrightarrow$ *to express that* $\to$ *is a* $\twoheadrightarrow$*-strategy, i.e. part of* $\twoheadrightarrow$*;*
- $\leftarrow$ *converse of notation* $\to$ *to expresses that* $\leftarrow$ *is converse of* $\to$*; and*
- ${\color{red}\twoheadleftarrow}\!{\color{blue}\rightarrowtriangle}$ *union of notations* ${\color{red}\twoheadleftarrow}$ *and* ${\color{blue}\rightarrowtriangle}$ *to express* ${\color{red}\twoheadleftarrow}\!{\color{blue}\rightarrowtriangle}$ *is union of* ${\color{red}\twoheadleftarrow}$ *and* ${\color{blue}\rightarrowtriangle}$*.*

We use sub/superscripts to express restricting/extending the corresponding notion.

At the core of this paper are the following *commutation* versions for pairs of rewrite systems of standard notions for single rewrite systems (PN is the obvious 'peak' version of NF).

▶ **Definition 3.** We say the *normal form* property (NF) holds if $a \; {\color{red}\twoheadleftarrow}\!{\color{blue}\rightarrowtriangle}^* \; b$ with a in ${\color{blue}\rightarrowtriangle}$-normal form implies $a \; {}^*{\color{red}\twoheadleftarrow} \; b$, the *peak normal form* property (PN) holds if $a \; {}^*{\color{red}\twoheadleftarrow} \cdot {\color{blue}\rightarrowtriangle}^* \; b$ with a in ${\color{blue}\rightarrowtriangle}$-normal form implies $a \; {}^*{\color{red}\twoheadleftarrow} \; b$, the *Church–Rosser* property (CR) holds if $a \; {\color{red}\twoheadleftarrow}\!{\color{blue}\rightarrowtriangle}^* \; b$ implies $a \; {\color{blue}\rightarrowtriangle}^* \cdot {}^*{\color{red}\twoheadleftarrow} \; b$, and *commutation* (CO) holds if $a \; {}^*{\color{red}\twoheadleftarrow} \cdot {\color{blue}\rightarrowtriangle}^* \; b$ implies $a \; {\color{blue}\rightarrowtriangle}^* \cdot {}^*{\color{red}\twoheadleftarrow} \; b$.

CR is also known as e.g. factorisation, postponement, preponement, and separation, cf. [4]. Instantiating both ${\color{red}\twoheadrightarrow}$ and ${\color{blue}\rightarrowtriangle}$ to $\to$ yields the usual notions NF, PN, CR and CO (then called *confluence*). That NF $\iff$ PN $\Longleftarrow$ CR $\iff$ CO is folklore, cf. [4]. In order to be able to express that one rewrite system is 'better' than another, we enrich these notions/diagrams in two ways: we equip steps with a measure (with respect to which conversions are compared) and we allow for infinite reductions (defining all such to be 'worse' than finite ones).

▶ **Definition 4.** Let $\mathcal{M}$ be a monoid $(M, +, \bot)$. We call $\twoheadrightarrow$ an $\mathcal{M}$*-measured* rewrite system if it comes equipped with a map from $\twoheadrightarrow$-steps to $M - \{\bot\}$. The (*derivation*) *measure* of a sequence of $\twoheadrightarrow$-steps is the sum of the measures of its steps, from left to right. The (*conversion*) *measure* of a sequence of $\twoheadleftarrow$- and $\twoheadrightarrow$-steps is a pair having as first component the sum of the measures of its $\twoheadleftarrow$-steps, from right to left, and as second component the sum of the measures of its $\twoheadrightarrow$-steps, from left to right. $\mathcal{M}$ is a *derivation* monoid if it comes

[1] Only for systems having at most one step between any pair of objects do we speak of rewrite *relations*.
[2] Sub-rewrite systems that do change the set of normal forms, e.g. call-by-value for the $\lambda\beta$-calculus, are *not* strategies (they result in different *calculi*).

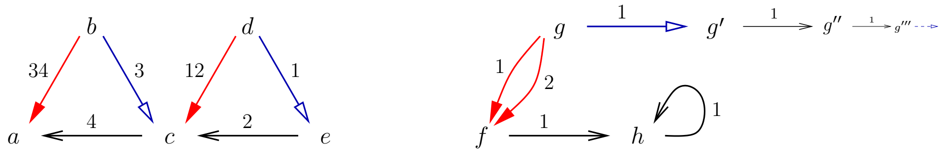

Figure 1 Measured rewrite systems on monoid of strings of digits ordered by embedding. Here $\rightarrow$ is used to denote steps in belonging to both $\rightarrow$ and $\rightarrow$, with the same measure.

equipped with a well-founded partial order $\leq$ such that $\bot$ is the *least* element ($\bot \leq m$ for all m) and $+$ is *strictly monotonic* in both arguments (if $m < n$ then $m + k < n + k$ and $k + m < k + n$, for all m, n, k). A derivation monoid is *cancellative* if $m + k \geq n + k$ entails $m \geq n$ for all m, n, k. A *conversion* monoid is a commutative derivation monoid.

Excluding $\bot$ as measure of steps will ensure that by prefixing/suffixing a step to a reduction its measure strictly increases.We use subscripts to indicate measures; in Fig. 1 digit-strings.

▶ **Example 5.** In the measured rewrite system $\rightarrow$ that is the union of $\rightarrow$ and $\rightarrow$ in Fig. 1, we have $a\ {}_{1234}\!\leftrightarrow^*_{31}\ e$ as witnessed by the conversion $a\ {}_{34}\!\leftarrow b \rightarrow_3 c\ {}_{12}\!\leftarrow d \rightarrow_1 e$, with the first component 1234 of the measure obtained by concatenating the measures 12 and 34 of the $\leftarrow$-steps, and the second component 31 by the measures 3 and 1 of the $\rightarrow$-steps. In general, strings with concatenation and empty string ordered by embedding constitute a cancellative but non-commutative derivation monoid. Assuming commutativity yields the cancellative conversion monoid of multisets with multiset sum and empty multiset. More generally, equipping a string/multiset monoid with the order generated by the union of embedding with some simply terminating string/multiset rewrite system yields a derivation/conversion monoid which need *not* be cancellative; consider the orders generated by $ab \to ac/[a, b] \to [a, c]$.

▶ **Example 6.** Measuring all steps of a rewrite system by 1 in the conversion monoid of the natural numbers with addition and zero equipped with less–than–or-equal, the measure μ of a reduction $\rightarrow^*_\mu$ corresponds to its *length*, and for ${}_n\!\leftrightarrow^*_m$ the integer $m - n$ corresponds to the *difference* of the conversion, the number of $\rightarrow$-steps minus the number of $\leftarrow$-steps [20, 21].

Next, we consider ordering [16, Definition 5] and pasting [17, Examples 4,9] diagrams.

▶ **Definition 7.** A conversion is *ordered* if $m \geq n$ for its measure (m, n), and *cyclic* if its source and target are the same. *Shift* equivalence is generated by identifying a cyclic conversion $C \cdot D$ with $D \cdot C$. A (*conversion*) *diagram* is a pair C, D of conversions with the same sources and targets. Its *induced* conversion is $C^{-1} \cdot D$, inducing notions of measure and shift equivalence on diagrams, cf. [24]. *Pasting* diagrams shift equivalent to C, D and D, E (*on* D) gives C, E.

Occurrences of D, determined by picking an object on the cyclic conversion (the source of D) and a length (of D), may be used to disambiguate pasting. For commutative monoids shift equivalent diagrams have the same measure. A *derivation* diagram comprises two reductions.

▶ **Example 8.** In Fig. 1, we have, e.g., two diagrams $b \rightarrow_{34} a, b \rightarrow_3 c \rightarrow_4 a$ and $d \rightarrow_{12} c, d \rightarrow_1 e \rightarrow_2 c$ having underlying conversions $a\ {}_{34}\!\leftarrow b \rightarrow_3 c \rightarrow_4 a$ respectively $c\ {}_{12}\!\leftarrow d \rightarrow_1 e \rightarrow_2 c$. The former is shift equivalent to the diagram $c\ {}_3\!\leftarrow b \rightarrow_{34} a\ {}_4\!\leftarrow c, c$ the latter to $c, c\ {}_{12}\!\leftarrow d \rightarrow_1 e \rightarrow_2 c$, and pasting them on c yields $c\ {}_3\!\leftarrow b \rightarrow_{34} a\ {}_4\!\leftarrow c, c\ {}_{12}\!\leftarrow d \rightarrow_1 e \rightarrow_2 c$. Whereas the first of the original diagrams is an ordered derivation diagram (both the 'counterclockwise' and 'clockwise' measures of the 1st are 34) its shift equivalent is not ($34 \not\geq 43$), and the result of pasting is neither shift equivalent to a derivation diagram nor to an ordered diagram.

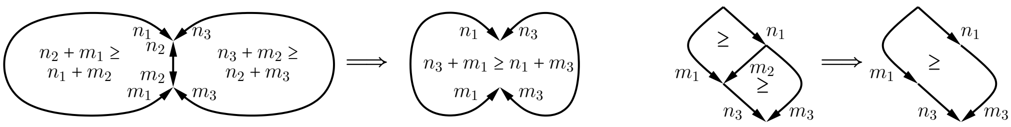

Figure 2 Preservation of order ('counterclockwise' $\geq$ 'clockwise') by pasting on ${}_{n_2}\!\leftrightarrow^*_{m_2}$.

Pasting derivation diagrams on reductions requires associativity of $+$, for conversion diagrams it also requires commutativity, and for pasting on conversions also cancellation is needed:

▶ **Lemma 9.**

- *Pasting preserves order for cancellative conversion monoids (Fig. 2 left);*
- *Pasting on a reduction (m_2 or n_2 is $\perp$ in Fig. 2) preserves order for conversion monoids;*
- *Pasting ordered derivation diagrams on a reduction with the source of one and target of the other, gives a diagram shift equivalent to an ordered derivation diagram (Fig. 2 right).*

Proof. Let conversions C_i have measure (n_i, m_i) in the pasted diagrams C_1, C_2 and C_2, C_3. In the first two items, orderedness yields $n_2 + m_1 \geq n_1 + m_2$ and $n_3 + m_2 \geq n_2 + m_3$.

- Combining both yields $n_2 + m_1 + n_3 + m_2 \geq n_1 + m_2 + n_2 + m_3$ from which we conclude to $n_3 + m_1 \geq n_1 + m_3$ by commutativity and cancelling $m_2 + n_2$;
- If w.l.o.g. $n_2 = \perp$, the assumptions yield $n_3 + m_1 \geq n_3 + n_1 + m_2 \geq n_1 + m_3$ by commutativity;
- If w.l.o.g. the reduction has the same target as the original 1st diagram and the same source as the 2nd (see Fig. 2 right), then orderedness of those original derivation diagrams yields $m_1 \geq n_1 + m_2$ and $m_2 + n_3 \geq m_3$, so $m_1 + n_3 \geq n_1 + m_2 + n_3 \geq n_1 + m_3$. ◀

▶ **Definition 10.** The rewrite system $\rightarrowtriangle^{\infty}$ [4] has the same objects as $\rightarrowtriangle$ and a step from a to b for each infinite $\rightarrowtriangle$-rewrite sequence from a and for any b. Given a monoid $\mathcal{M}$, the monoid $\mathcal{M}^{\top}$ is obtained by adjoining a fresh element $\top$ to the carrier and defining $\top + \top = \top + m = m + \top = \top$ for all $m \in M$. Given a $\mathcal{M}$-measured rewrite system $\rightarrowtriangle$, mapping $\rightarrowtriangle^{\infty}$-steps to $\top$ gives rise to an $\mathcal{M}^{\top}$-measured rewrite system $\rightarrowtriangle \cup \rightarrowtriangle^{\infty}$, and similarly for a pair $\to, \rightarrowtriangle$ (see Definition 4).

Although $\rightarrowtriangle^{\infty}$ is not common in rewriting yet, it is in relational program semantics [4]. The extension from $\mathcal{M}$ to $\mathcal{M}^{\top}$ preserves commutativity, and cancelling elements of $\mathcal{M}$. (Although not needed here, note the first item of Lemma 9 needs m_2 or n_2 to be finite to go through). Two rewrite systems particularly important for this paper are $(\rightarrowtriangle \cup \rightarrowtriangle^{\infty})^*$ and $(\leftarrow \cup \rightarrowtriangle \cup \rightarrowtriangle^{\infty})^*$ which we will refer to as *extended* reduction ($\rightarrowtriangle^{\circledast}$) and conversion ($\leftrightarrowtriangle^{\circledast}$), respectively. Beware that locations of superscipts matter, e.g. $\leftarrowtriangle^{\infty}$ would be distinct from the converse of $\rightarrowtriangle^{\infty}$.

▶ **Example 11.** In the measured rewrite system of Fig. 1 $f\ {}_{2}\!\leftrightarrowtriangle^{\circledast}_{\top}\ b$ since $f\ {}_{2}\!\leftarrow g \rightarrowtriangle^{\infty}_{\top} b$ as g admits an infinite $\rightarrowtriangle$-reduction $g \rightarrowtriangle g' \rightarrowtriangle \ldots$. We do *not* have $b \leftrightarrowtriangle^{\circledast} f$.

We may assume $\rightarrowtriangle^{\infty}$ to occur only, if at all, at the end of an extended reduction or conversion because $\rightarrowtriangle^{\circledast} = \rightarrowtriangle^* \cup \rightarrowtriangle^{\infty} = \rightarrowtriangle^* \cdot (\rightarrowtriangle^{\infty})^{=}$ and $\leftrightarrowtriangle^{\circledast} = \leftrightarrowtriangle^* \cdot (\rightarrowtriangle^{\infty})^{=}$, with superscript $=$ denoting reflexive closure. This process does not increase the first component of the measure and leaves the second unchanged. To differentiate between elements of $\mathcal{M}$ and $\mathcal{M}^{\top}$, we henceforth use $m, n, k, \ldots$ to range over the former and $\mu, \nu, \kappa, \ldots$ to range over the latter. We call the former *finite* as can be vindicated by setting infinite sums of non-$\perp$-elements to $\top$ and noting that infinite measures are not affected by (un)folding $\rightarrowtriangle^{\infty}$: if $a \rightarrowtriangle^{\infty}_{\top} b$ is witnessed by $a \rightarrowtriangle_m a' \rightarrowtriangle_{m'} a'' \rightarrowtriangle_{m''} \ldots$ then its measure is $m + m' + m' + \ldots = \top$, and so is the measure of $a \rightarrowtriangle_m a' \rightarrowtriangle^{\infty}_{\top} b$ because $m + \top = \top$. Thus, a reduction is infinite if and only if the corresponding extended reduction has measure $\top$ with respect to the length measure (Example 6).

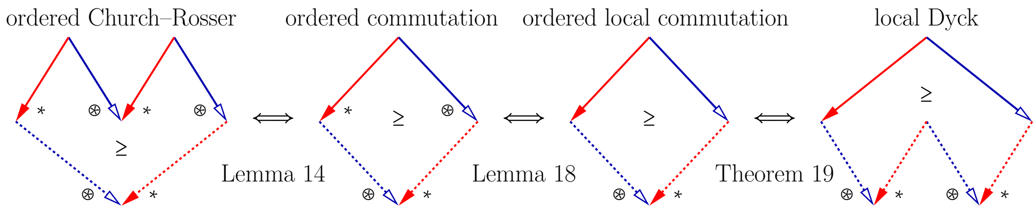

Figure 3 Localising ordered Church–Rosser, restricting the $\forall$, widening the $\exists$, to local Dyck.

3 Ordered commutation and Dyck diagrams

We introduce a property, *ordered* Church–Rosser, sufficient for the measure of the $\to$-reductions from a given object to be an *upper bound* on the measures of its $\rightarrowtriangle$-reductions. This will be used in Sections 5–6 to show normalisation of the latter via that of the former. Here we work towards deciding the property, *localising* it to the *local Dyck* property, Fig. 3.

Throughout, we assume $\to, \rightarrowtriangle$ are measured rewrite systems for the same derivation monoid.

▶ **Definition 12.** We say the *ordered* normal form property (ONF) holds if $a \;{}_{n}{\leftrightarrow}^{\circledast}_{\mu}\; b$ with a in $\rightarrowtriangle$-normal form implies $a \;{}_{n'}^{\;*}{\leftarrow}\; b$, the *ordered* peak normal form property (OPN) holds if $a \;{}_{n}^{\;*}{\leftarrow} \cdot \rightarrowtriangle^{\circledast}_{\mu}\; b$ with a in $\rightarrowtriangle$-normal form implies $a \;{}_{n'}^{\;*}{\leftarrow}\; b$, the *ordered* Church–Rosser property (OCR) holds if $a \;{}_{n}{\leftrightarrow}^{\circledast}_{\mu}\; b$ implies $a \rightarrowtriangle^{*}_{\mu'} \cdot {}_{n'}^{\;*}{\leftarrow}\; b$, and *ordered* commutation (OCO) holds if $a \;{}_{n}^{\;*}{\leftarrow} \cdot \rightarrowtriangle^{\circledast}_{\mu}\; b$ implies $a \rightarrowtriangle^{\circledast}_{\mu'} \cdot {}_{n'}^{\;*}{\leftarrow}\; b$, with the *order* constraint $n + \mu' \geq \mu + n'$ ($\mu' = \bot$ by default).

In words, a commutation diagram is ordered if the measure of its left leg is as large as that of its right leg. Note that this corresponds exactly to orderedness (see Definition 7) of the corresponding derivation diagram for derivation monoids. Similarly, for OCR the constraint corresponds to orderedness of the corresponding conversion diagram for conversion monoids. Remark that if OCR/OCO and $a \leftrightarrow^{\circledast} b$/ $a \;{}^{*}{\leftarrow} \cdot \rightarrowtriangle^{\circledast} b$, then if a is $\rightarrowtriangle$-terminating so is b and the corresponding OCR, OCO diagram is finite. Lemma 9 vindicates pasting such diagrams. That OCO need *not* imply commutation or confluence, follows by considering non $\rightarrowtriangle$-terminating such a; OCO holds in Example 13, but commutation not for $f \;{}_{i}{\leftarrow}\; g \rightarrowtriangle_{1} g'$.[3]

▶ **Example 13.** In Fig. 1, OCO and OPN are easily seen to hold by considering the three 'interesting' local peaks $a \;{}_{34}{\leftarrow}\; b \rightarrowtriangle_{3} c$, $c \;{}_{12}{\leftarrow}\; d \rightarrowtriangle_{1} e$ and $f \;{}_{i}{\leftarrow}\; g \rightarrowtriangle_{1} g'$, that are completed into ordered commutation diagrams by respectively $a \;{}_{4}{\leftarrow}\; c$, $c \;{}_{2}{\leftarrow}\; e$ and $f \rightarrowtriangle^{\infty}_{\top} g'$.

Composing the first two of these local peaks (on c) yields a conversion $a \;{}_{1234}{\leftrightarrow}^{*}_{31}\; e$ that can (only) be completed into a commutation diagram by $a \;{}_{24}^{\;*}{\leftarrow}\; e$, which does not satisfy the order constraint as $1234 \not\geq 3124$, showing neither OCR nor ONF holds, cf. Example 8.

The four notions relate to each other as in the unordered (finite) case.

▶ **Lemma 14.** *ONF $\iff$ OPN $\impliedby$ OCR $\iff$ OCO, for conversion monoids.*

Proof. All implications hold by definition except for OCO $\implies$ OCR and OPN $\implies$ ONF. These are shown by induction on the number of peaks in a conversion, pasting diagrams as in the unordered case [4]. Since the diagrams are conversion diagrams, a conversion monoid is needed (cf. Example 13) to let order be preserved by pasting on reductions (Lemma 9). ◀

[3] The converse also fails as witnessed by $b_0 \rightarrowtriangle b_1 \rightarrowtriangle \ldots$ and $a \leftarrow b_i$ for all i.

Note that the four notions are asymmetric in that they consider infinite $\rightarrowtriangle$-reductions but only finite $\to$-reductions. However, this sufficies for *bounding* $\rightarrowtriangle$-reductions by $\to$-reductions:

▶ **Proposition 15.** *If OPN and $\to, \rightarrowtriangle$ have the same normal forms, then for any peak $a \;{}_{\nu}^{\circledast}\!\!\leftarrow \cdot \rightarrowtriangle_{\mu}^{\circledast} b$ of maximal reductions, $\nu \geq \mu$ and if the left leg is finite so is the right leg and $a = b$.*

Proof. If the left leg of the peak has infinite length, then ν is $\top$ and we conclude trivially. Otherwise, the peak has shape $a \;{}_{n}^{*}\!\!\leftarrow \cdot \rightarrowtriangle_{\mu}^{\circledast} b$ for some n with a in normal form. By OPN for it, there exists $a \;{}_{m'}^{*}\!\!\leftarrow b$ such that $n \geq \mu + m'$. As n is finite, so is μ, so the right leg of the peak is finite and by maximality must end in a normal form and $a = b$. ◀

In particular, if $\to, \rightarrowtriangle$ are strategies for the same rewrite system, as is, e.g., the case for the systems in Fig. 1, normalisation of the former entails normalisation of the latter.[4]

Having shown their usefulness, we turn to localising the properties. *Localisation* of a ∀∃-property aims at finding an equivalent property that *restricts* the domain of the ∀-quantifier and *widens* that of the ∃-quantifier, to enable or ease deciding it (automatically). The classical example is localisation of the Church–Rosser property (∀conversions ∃valley) by *restricting* first to peaks, then further to local peaks [13] and finally by *widening* to conversions below the source [22], for terminating rewrite systems. Here, as we already have OCR $\iff$ OCO, we restrict OCO to ordered local commutation and widen that to the local Dyck property.

▶ **Definition 16.** Ordered *local* commutation arises from OCO by restricting both legs of the peak to reductions of length 1. A diagram comprising a local peak $a \;{}_{n}\!\!\leftarrow \cdot \rightarrowtriangle_{m} b$ and extended conversion $a \;{}_{n'}\!\!\leftrightarrow_{\mu'}^{\circledast} b$ is a *Dyck* diagram if $n + \mu' \geq m + n'$ and the *Dyck*-condition holds: for every prefix (of which there are finitely many) $a \;{}_{n''}\!\!\leftrightarrow_{m''}^{\circledast} c$ of the conversion $n + m'' > n''$. We say the systems are *locally Dyck* if each such peak can be completed into a Dyck diagram.

Our naming is based on that for the length measure the number of backward ($\leftarrow$) steps in the conversion then never exceeds the number of forward ($\rightarrowtriangle$) steps, as in the Dyck language.

▶ **Example 17.** Let for some N, the abstract rewrite system $\to$ be given by $b_i \to b_{i+1} \leftarrow a_i \to c_{i+1} \leftarrow c_i$ for all $1 \leq i \leq N$ and $b_{N+1} \to c_{N+1}$, and $\rightarrowtriangle$ be the $\to$-strategy comprising all $\to$-steps except those from a_i to b_{i+1}, both with respect to the length measure. The only interesting local peaks are $b_{i+1} \leftarrow a_i \rightarrowtriangle c_{i+1}$ which can be completed into a Dyck diagram; for $i < N$ by $b_{i+1} \rightarrowtriangle b_{i+2} \leftarrow a_{i+1} \rightarrowtriangle c_{i+2} \leftarrow c_{i+1}$ (the conditions for its 5 prefixes are respectively $1 > 0$, $2 > 0$, $2 > 1$, $3 > 1$ and $3 > 2$) and for $i = N$ by $b_{N+1} \rightarrowtriangle c_{N+1}$. For the system, this takes a number of conversion (back-and-forth) steps linear in N, whereas naïvely completing into ordered commutation diagrams requires a quadratic number of reduction steps, cf. [17, Example 8]. Hence, using the following, the length of $\to$-reductions bounds those of $\rightarrowtriangle$-reductions.

▶ **Lemma 18.** *Ordered commutation iff ordered local commutation.*

Proof. The only–if-direction is trivial. For the if-direction, it suffices to consider peaks $a \;{}_{\hat{n}}^{*}\!\!\leftarrow \cdot \rightarrowtriangle_{\hat{\mu}}^{\circledast} b$ such that a is $\rightarrowtriangle$-terminating as otherwise we conclude by $a \rightarrowtriangle_{\top}^{\infty} b$. We show such a peak can be completed by $a \rightarrowtriangle_{\mu'}^{\circledast} \cdot \;{}_{n'}^{*}\!\!\leftarrow b$ with $\hat{n} + \mu' \geq \hat{\mu} + n'$ into an OCO diagram, finite by the Remark above Example 13, by induction on $(a, \hat{n})$ ordered by the *leg* order

[4] For such strategies the proposition shows that OPN is *sufficient* for $\rightarrowtriangle$ being universally better than $\to$, in the sense of [16], and hence [16] that $\rightarrowtriangle$ is normalising, minimal (if $\rightarrowtriangle \subseteq \to$) and $\to$ is perpetual, maximal (if $\to \subseteq \rightarrowtriangle$). OPN is not *necessary* for it: $b \leftarrow a \rightarrowtriangle c$.

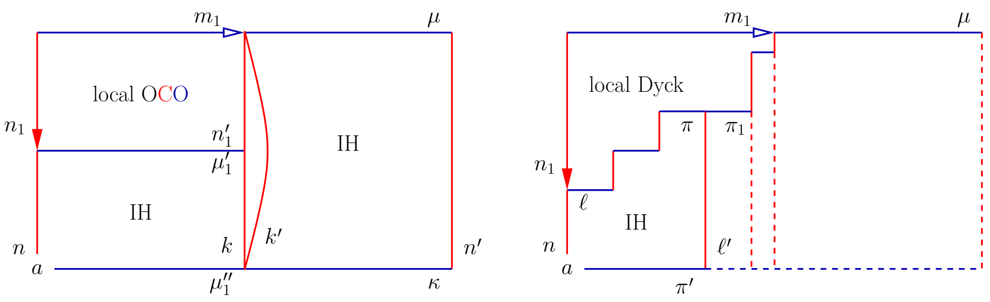

Figure 4 Proofs of Lemma 18 (left) and Theorem 19 (right).

$>_\bullet$, the lexicographic product of $\to^{+}$ and $>$. If either leg of the peak is empty, it is trivial. Otherwise, it has shape $a\ {}_{n}^{*}\!\leftarrow a_1\ {}_{n_1}\!\leftarrow \cdot \to_{m_1} b_1 \to_{\mu}^{\circledast} b$ with $\hat{n} = n_1 + n$ and $\hat{m} = m_1 + \mu$; Fig. 4.

⟦ Ordered local commutation applied to $a_1\ {}_{n_1}\!\leftarrow \cdot \to_{m_1} b_1$ yields $a_1 \to_{\mu_1'}^{\circledast} c_1\ {}_{n_1'}^{*}\!\leftarrow b_1$ with $n_1 + \mu_1' \geq m_1 + n_1'$. Consider the peak $a\ {}_{n}^{*}\!\leftarrow a_1 \to_{\mu_1'}^{\circledast} c_1$. The induction hypothesis applies to it as $\hat{n} = n_1 + n > n$, giving $a \to_{\mu_1''}^{\circledast} a'\ {}_{k}^{*}\!\leftarrow c_1$ with $n + \mu_1'' \geq \mu_1' + k$. Vertically pasting this diagram to the one for the local peak yields a diagram satisfying $n_1 + n + \mu_1'' \geq n_1 + \mu_1' + k \geq m_1 + n_1' + k$, i.e. that is ordered. Abbreviating $n_1' + k$ to k' it has a valley $a \to_{\mu_1''}^{\circledast} a'\ {}_{k'}^{*}\!\leftarrow b_1$. ⟧

Now consider the peak $a'\ {}_{k'}^{*}\!\leftarrow b_1 \to_{\mu}^{\circledast} b$. The induction hypothesis applies to it as either $a \to_{\mu_1''}^{*} a'$ is not empty, or it is empty and then $a = a'$, μ_1'' is $\perp$, and instantiating the inequality above yields $n_1 + n \geq m_1 + k' > k'$. Hence we obtain a valley $a' \to_{\kappa}^{\circledast} \cdot\ {}_{n'}^{*}\!\leftarrow b$ with $k' + \kappa \geq \mu + n'$. Setting μ' to $\mu_1'' + \kappa$ we conclude by horizontal diagram pasting, yielding the order constraint $n_1 + n + \mu' = n_1 + n + \mu_1'' + \kappa \geq m_1 + k' + \kappa \geq m_1 + \mu + n'$. ◀

▶ **Theorem 19.** *Ordered commutation iff locally Dyck, for cancellative conversion monoids.*

Proof. For the only–if-direction it suffices to remark that in an ordered local commutation diagram $\to$-steps precede $\leftarrow$-steps, so that orderedness entails the Dyck-condition. For the if-direction, we proceed exactly as in the proof of Lemma 18 but using the local Dyck property instead of ordered local commutation. That is, we replace the part between ⟦ and ⟧ there by:

The local Dyck property applied to $a_1\ {}_{n_1}\!\leftarrow \cdot \to_{m_1} b_1$ yields $a_1\ {}_{n_1'}^{*}\!\leftrightarrow_{\mu_1'}^{\circledast} b_1$ such that $n_1 + \mu_1' \geq m_1 + n_1'$ and satisfying the Dyck-condition: for every prefix $a_1\ {}_{\ell}^{*}\!\leftrightarrow_{\pi}^{\circledast} c$ of the conversion, $n_1 + \pi > \ell$. We show by a sub-induction on the length of the prefix, that there exists a valley $a \to_{\pi'}^{\circledast} d\ {}_{\ell'}^{*}\!\leftarrow c$ completing $a\ {}_{n}^{*}\!\leftarrow a_1\ {}_{\ell}^{*}\!\leftrightarrow_{\pi}^{\circledast} c$ into an ordered diagram, i.e. such that $\ell + n + \pi' \geq \pi + \ell'$. The case of the empty prefix being trivial, assume the property holds for a given prefix up to c, and distinguish cases on the next step of the prefix.

If $c\ {}_{\ell_1}\!\leftarrow c_1$, then we simply affix it: $a\ {}_{n}^{*}\!\leftarrow a_1\ {}_{\ell}^{*}\!\leftrightarrow_{\pi}^{\circledast} c\ {}_{\ell_1}\!\leftarrow c_1$ is completed by the valley $a \to_{\pi'}^{\circledast} d\ {}_{\ell'}^{*}\!\leftarrow c\ {}_{\ell_1}\!\leftarrow c_1$ into an ordered diagram: $\ell_1 + \ell + n + \pi' \geq \ell_1 + \pi + \ell'$.

If $c \to_{\pi_1} c_1$ or $c \to_{\pi_1}^{\infty} c_1$, then consider the peak between it and $d\ {}_{\ell'}^{*}\!\leftarrow c$, see Fig. 4. That the main induction applies to it follows (by a decrease in the 1st component) in case π' is not $\perp$ and otherwise (by a decrease in the 2nd component) by combining orderedness with the Dyck condition: $n_1 + \ell + n \geq n_1 + \pi + \ell' > \ell + \ell'$ hence $n_1 + n > \ell'$ by cancelling ℓ. Thus we obtain a valley $d \to_{\pi_1'}^{\circledast} d_1\ {}_{\ell'}^{*}\!\leftarrow c_1$ completing the peak to an ordered diagram. (By the Remark above Example 13, this shows $c \to_{\pi_1}^{\infty} c_1$ is in fact impossible.) Pasting it to the one of the IH yields an ordered diagram with valley $a \to_{\pi'}^{\circledast} d \to_{\pi_1'}^{\circledast} d_1\ {}_{\ell'}^{*}\!\leftarrow c_1$, as desired.

Let $a \to_{\mu_1''}^{\circledast} a'\ {}_{k'}^{*}\!\leftarrow b_1$ be the valley thus obtained for the whole of $a_1\ {}_{n_1'}^{*}\!\leftrightarrow_{\mu_1'}^{\circledast} b_1$. By the induction hypothesis it satisfies the order constraint $n_1' + n + \mu_1'' \geq \mu_1' + k'$. Combining it with the order constraint $n_1 + \mu'_1 \geq m_1 + n'_1$ of the local Dyck diagram, adding the respective sides and cancelling $n'_1 + \mu'_1$ gives $n_1 + n + \mu''_1 \geq m_1 + k'$, showing that the valley completes the peak $a \; {}_{n}^{*}\!\!\leftarrow a_1 \; {}_{n_1}\!\!\leftarrow \cdot \rightarrow_{m_1} b_1$ into an ordered commutation diagram. ◀

For non-cancellative conversion monoids the if-direction may fail. The theorem allows one to *localise* showing that a strategy $\rightarrow$ bounds another strategy $\rightarrow$, by checking all *local* peaks between both to be completable into Dyck diagrams. We give 2 typical examples, cf. [16].

▶ **Example 20** (‡). For the λ-calculus with η-reduction $\rightarrow$ (see Example 40 for the rewrite rule), consider the innermost strategy $\rightarrow$, and, based on the idea that checking applicability of the η-rule involves checking absence of the bound variable from the body, measure a step by the *size* of its body. By orthogonality, linearity and preservation of innermost redexes, a local peak $N \; {}_{n}\!\!\leftarrow M \rightarrow_m P$ either is trivial ($N = P$) or it can be completed by a valley of shape $N \rightarrow_{m'} Q \; {}_{n'}\!\!\leftarrow P$. If the redexes are disjoint, then $m' = m$ and $n' = n$, so the diagram is Dyck. Otherwise, the latter is in the body of the former and we conclude to $n + m' \geq m + n'$ again because $m' = m$ and $n' = n - 2$ (an @ and λ have disappeared). Thus $\rightarrow$ is optimal, giving a lower bound on the (size) measure of η-reduction.

▶ **Example 21** (‡). For the λ-calculus with β-reduction, let $\rightarrow$ be the leftmost–outermost strategy, $\rightarrow$ be the needed strategy [3], and consider a (non-trivial) local peak $N \leftarrow M \rightarrow P$. By orthogonality and projection of leftmost–outermost steps over other steps $N \rightarrow_{\beta}^{*} Q \; {}_{\beta}\!\!\leftarrow P$ with the former a development of the residuals of $M \rightarrow P$. Factorising the former into needed steps followed by non-needed steps, and observing that the latter is a leftmost–outermost step again $N \rightarrow^{+} N' \rightarrow_{\beta}^{*} Q \leftarrow P$ with the first non-empty by definition of neededness. By repeatedly contracting a leftmost–outermost redex starting from N' and performing its projection on Q until (if this happens at all) the terms reached by both are the same, yields a $\rightarrow$-reduction from N' (using leftmost–outermost redexes are needed) and a $\rightarrow$-reduction from Q of the same (possibly infinite) length. Therefore, $N \rightarrow_{\geq 1} N' \rightarrow_{\mu}^{\circledast} Q' \; {}_{\mu}^{\circledast}\!\!\leftarrow Q \; {}_{1}\!\!\leftarrow P$, i.e. a Dyck diagram (note that if μ is ∞, then $N' \rightarrow_{\infty}^{\circledast} Q$). Thus $\rightarrow$ is pessimal, giving an upper bound on the (length) measure of needed β-reduction (but not on non-needed; cf. $(\lambda x.y)\Omega$).

4 Ordered confluence and random descent

We show that instantiating both $\rightarrow, \rightarrow$ to the same rewrite system $\rightarrow$, and assuming it to be measured by a *conversion* monoid, all conditions of the previous section are equivalent, and sufficient to conclude that $\rightarrow$ *bounds itself*. We show that the bounding system ($\rightarrow$) being the same as the system being bound ($\rightarrow$) allows to replace the *order* constraint in the ordered normal form property by *equality*, yielding Newman's *random descent* property expressing that "if an end-form exists it is reached by random descent" [13, 20, 21, 16]. We localise it to the local Dyck property and give several examples.

▶ **Definition 22.** $\rightarrow$ has *random descent* (RD) [16] if $a \; {}_{n}\!\!\leftrightarrow_{\mu}^{\circledast} b$ with a in normal form, implies $a \; {}_{n'}^{*}\!\!\leftarrow b$ with $n = \mu + n'$. *Peak random descent* (PR) is obtained by restricting to $a \; {}_{n}^{*}\!\!\leftarrow \cdot \rightarrow_{\mu}^{\circledast} b$.

RD and PR being more strict versions of NF respectively PN, they have similar properties. In particular, RD $\implies$ PR, and if $a \; {}^{*}\!\!\leftarrow b$ with a in normalform and either RD or PR holds, then b is terminating. Note the n' in the definition only depends on b and is unique since applying either property to $a \; {}_{n'}^{*}\!\!\leftarrow b \rightarrow_{n''}^{*} a'$ for normal forms a, a' gives $a = a'$ and $n' = n''$. This justifies defining the *distance* $\mathsf{d}(b)$ of such an object b convertible to normal form to be that n', setting $\mathsf{d}(b)$ to ∞ for objects not convertible to normal form.

▶ **Example 23.** β-reduction does not have the random descent property for the length measure as witnessed e.g. by $(\lambda xz.zxx)((\lambda x.x)y)$ which allows reductions to normal form $\lambda z.zyy$ of lengths both 2 and 3. The restriction of β to the leftmost–outermost strategy does have RD because it is deterministic. The restriction to *linear* β-redexes, i.e. such that the bound variable occurs exactly once in the body, has RD because redex-patterns cannot be replicated.[5] On its own, η-reduction has RD as it is linear and orthogonal.

▶ **Lemma 24.** *ONF $\iff$ OPN $\iff$ OCR $\iff$ OCO $\iff$ PR $\iff$ RD, all are equivalent to ordered local confluence, and for cancellative monoids to the local Dyck property.*

Proof. By Lemma 14 and the above, to conclude to the implications on the first line, it suffices to show ONF $\implies$ RD and OCO $\impliedby$ PR. For the former, suppose $a \;{}_n{\leftrightarrow}^{\circledast}_{\mu}\; b$ with a in normal form. By ONF, $a \;{}_{n'}^{*}{\leftarrow}\; b$ with $n \geq \mu + n'$, so μ is finite. By ONF for $a \;{}_{n'}^{*}{\leftarrow}\; b \;{}_{\mu}{\leftrightarrow}^{*}_{n}\; a$, also $\mu + n' \geq n$. For the latter, consider a peak $a \;{}_{n}^{*}{\leftarrow} \cdot \rightarrow^{\circledast}_{\mu}\; b$. If a allows an infinite reduction, we conclude. Otherwise, $c \;{}_{m'}^{*}{\leftarrow}\; a$ for some normal form c. By PR for $c \;{}_{m'}^{*}{\leftarrow}\; a \;{}_{n}^{*}{\leftarrow} \cdot \rightarrow^{\circledast}_{\mu}\; b$, we obtain $a \;{}_{n'}^{*}{\leftarrow}\; b$ with $n + m' = \mu + n'$. The second line follows by Lemma 18 and Theorem 19. ◀

As in the previous section, the implication from a peak-property to its conversion-property fails for derivation monoids (Example 13), and the lemma allows to *localise* random descent. The examples here and in the previous section are illustrative both of localisation and of the flexibility offered by measuring by length, size of subterm or pattern, rule, ... Many possibilities come to mind with as extreme case measuring a reduction by 'itself', say as the string or multiset of its steps. Our final example uses the left–outer order on positions.

▶ **Example 25.** The single rule term rewrite system $f(x,x) \to f(x, f(x,a))$ has random descent (for the length measure) because there are no critical peaks and the rule is *variable preserving* in that all variables appear the same number of times in both the left- and right-hand sides. The latter condition guarantees that in the so-called variable-overlap case of the critical pair lemma, both legs of the resulting diagram have exactly the same length.

▶ **Example 26.** The term rewrite system given by

$$f(x,x) \;\to_1\; f(x,g(x)) \quad f(x,x) \;\to_2\; f(x,h(x)) \quad g(x) \;\to_1\; h(x) \quad c \;\to_1\; g(c)$$

has random descent with respect to the indicated *rule* measures (Definition 44, [17, Sect. 4.2]). As in Example 25 the rules are variable preserving, but now there is the critical peak (and its symmetric version) $f(x,g(x)) \leftarrow_1 f(x,x) \rightarrow_2 f(x,h(x))$ which however is completable by the step $f(x,g(x)) \rightarrow_1 f(x,h(x))$ into a Dyck diagram where both legs have measure 2.

▶ **Definition 27.** On positions in terms, the *left* relation is defined by $p \cdot i \cdot q \prec_l p \cdot j \cdot q'$ and the *outer* (or *prefix*) relation by $p \prec_o p \cdot i \cdot q$, for arbitrary positions p, q, q' and natural numbers $i < j$. The *left–outer* relation is defined by $\prec_{lo} = \prec_l \cup \prec_o$.

The relations $\prec_o$, $\prec_l$ and $\prec_{lo}$ are strict orders, $\prec_o$ and $\prec_l$ are disjoint, and $\preceq_{lo}$ is total.

▶ **Example 28** (‡)**.** The *spine* positions of a λ-term M are, if it has shape $\lambda\vec{x}.y\vec{M}$ then the displayed positions and the spine positions of the M_i prefixed by their position in M, and otherwise its *head spine* positions. The *head spine* positions of terms x, $\lambda x.M_1$, $M_1 M_2$ are all positions $\preceq_{lo}$-related to the position of M_1 and all head spine positions of M_1 prefixed by its position in M. A *spine* redex-pattern is a redex-pattern at spine position. Always contracting such we call a *spine* strategy, which is justified by the fact that any term not in $\beta(\eta)$-normal form has at least one spine $\beta(\eta)$-redex-pattern.[6] That the spine strategy has random descent

---

5 This is not a strategy for β-reduction in λ-calculus as e.g. $(\lambda xz.zxx)y$ is a normal form for it.

6 A λ-term not in η-normal form need not contain a spine η-redex-pattern, e.g. $(\lambda x.x)\lambda y.zy$.

for β-reduction [3, Proposition 4.21] follows from that the β-rule does not overlap with itself, and that contracting a spine β-redex-pattern leaves a unique descendant (residual) of any non-overlapping spine position (redex-pattern). As the η-rule does not overlap with itself and is variable preserving, descendants are unique after η-steps and η-reduction has random descent. Since the (two) critical peaks between the β- and η-rules are trivial, they are locally Dyck, from which it follows that the spine strategy has random descent for $\beta\eta$-reduction.

The underlying intuition, generalising that for externality [8], is that the spine is a prefix of a term that persists linearly (contrary to externality, it may involve change, but only 'linear change'). It applies to some other term rewrite systems, e.g. Combinatory Logic as well. We conclude by an easy result allowing to infer RD for strategies.

▶ **Lemma 29.** *If $\rightarrowtriangle$ has random descent then so does any strategy $\to$ for it.*

▶ **Example 30.** Since the spine strategy has random descent for $\beta\eta$-reduction (see above), so does the left–outer (see, the text above, Definition 41) strategy.

5 Compatibility

We present a method to establish hyper-normalisation of strategies for abstract rewrite systems, based on a diagram we (inspired by Staples' notion of *compatible* refinement, cf. [19, Exercise 1.3.9]) dub *compatibility* governing the interaction between the strategy and the system, and show it can be made flexible by well-foundedly indexing steps giving rise to the notion of *decreasing* compatibility. As an application, one may immediately conclude normalisation of the needed strategy for $\lambda\beta\eta$, cf. [9, Chapter IV] and [3], from that of the spine strategy (Example 38), using that the former is *bounded* by the latter (Example 21). Our methods rely on the strategy having the random descent property, for an arbitrary cancellative conversion monoid, as introduced above.

We assume $\to$ is a strategy having random descent for the rewrite system $\rightarrowtriangle$.

▶ **Definition 31.** $\to$ is (*ordered*) *compatible* with $\rightarrowtriangle$, if $a \rightarrowtriangle b$ entails $a \;{}_{n}{\leftrightarrow}^{\circledast}_{\mu}\; b$ (with $\mu \geq n$).

▶ **Example 32** (‡)**.** Spine reduction is compatible with backward $\beta(\eta)$-steps and ordered compatible with forward such steps. We provide a proof of this later, via decreasing compatibility, but the intuition for that it holds is that, given a $\beta(\eta)$-step $M \rightarrowtriangle N$ one may contract an arbitrary spine redex in M. In case this $\to$-step yields N then we are done. Otherwise, one may contract 'the same' spine redex in N, project $M \rightarrowtriangle N$ over both, and repeat the process until the first case applies. We also conclude when this process proceeds indefinitely, as then we have constructed infinite $\to$-reductions from M and N.

▶ **Lemma 33.** *$\to$ is (ordered) compatible with $\rightarrowtriangle$ iff it is (ordered) compatible with $\rightarrowtriangle^*$.*

Proof. The if-direction is trivial by $\rightarrowtriangle$ being contained in $\rightarrowtriangle^*$ and the only–if-direction follows by an easy induction on the length of $\rightarrowtriangle^*$-reductions (using commutativity). ◀

▶ **Theorem 34.** *If $\to$ is a random descent $\rightarrowtriangle$-strategy compatible with $\leftarrowtriangle$, it is $\rightarrowtriangle$-normalising. If, moreover, $\to$ is ordered compatible with $\rightarrowtriangle$, it is $\rightarrowtriangle$-hyper-normalising and $\rightarrowtriangle$ has NF.*

Proof. Suppose $\to$ is a random descent $\rightarrowtriangle$-strategy that is compatible with $\leftarrowtriangle$, and $a \rightarrowtriangle^* b$ with b in $\rightarrowtriangle$-normal form. Lemma 33 yields that $\to$ is compatible with $\leftarrowtriangle^*$, which applied to $b \leftarrowtriangle^* a$ entails $b \;{}_{n}{\leftrightarrow}^{\circledast}_{\mu}\; a$. By random descent a is $\to$-terminating.

Suppose, moreover, $\to$ is ordered compatible with $\rightarrowtriangle$, and a is $\rightarrowtriangle$-normalising. To prove $\to/\rightarrowtriangle$-reduction terminates on a, it suffices to show that for $a \rightarrowtriangle a'$ we have $\mathsf{d}(a) \geq \mathsf{d}(a')$ and if in fact $a \to a'$ then $\mathsf{d}(a) > \mathsf{d}(a')$. Distinguish cases on whether or not $a \rightarrowtriangle a'$ is a $\to$-step.

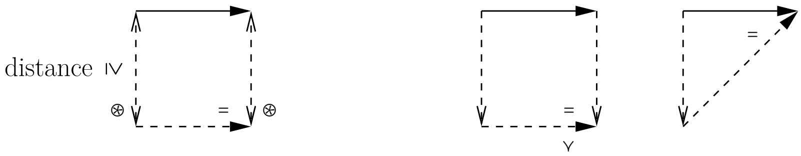

Figure 5 Decreasing compatibility (without condition, left) and its instance of Corollary 37.

- If $a \to_m a'$, then by RD for $b \leftarrow^*_{\mathsf{d}(a)} a \to_m a'$, $b \leftarrow^*_{\mathsf{d}(a')} a'$ with $\mathsf{d}(a) = m + \mathsf{d}(a') > \mathsf{d}(a')$.
- If $a \twoheadrightarrow a'$ then by $\to$ being ordered compatible with $\twoheadrightarrow$, we have $a \; {}_n{\leftrightarrow}^{\circledast}_{\mu} \; a'$ with $\mu \geq n$. By RD for $b \leftarrow^*_{\mathsf{d}(a)} a \; {}_n{\leftrightarrow}^{\circledast}_{\mu} \; a'$, then $b \leftarrow^*_{\mathsf{d}(a')} a'$ with $n + \mathsf{d}(a) = \mu + \mathsf{d}(a')$ (so μ is finite). Hence $\mathsf{d}(a) \geq \mathsf{d}(a')$ since in any cancellative derivation monoid if $n + k = m + \ell$ and $m \geq n$, then $n + k = m + \ell \geq n + \ell$ so by cancellation $k \geq \ell$ (without cancellation this property need not hold; let $\geq$ be generated by the multiset rule $[a, b] \to [a, a]$ and consider $[a, b, a]$).

Finally, we conclude to NF of $\twoheadrightarrow$ since the assumptions yield $\to$ is compatible with $\leftrightarrow$ hence with $\leftrightarrow^*$ by Lemma 33, and since $\to$ is a $\twoheadrightarrow$-strategy. ◀

By introducing a well-founded order on steps, as a parameter to the definition of compatibility, we increase its flexibility, cf. [20, Corollary 3.7], [21, Corollary 3]. The intuition captured is that $\twoheadrightarrow$-steps go 'at least as much forward (μ) as backward (n)' with respect to $\to$, recursively.

▶ **Definition 35.** We say $\to$ is (*ordered*) *decreasingly* compatible with $\twoheadrightarrow$, if for some well-founded order $\prec$ on the steps of $\twoheadrightarrow$, for all $\twoheadrightarrow$-steps ϕ, it holds $\twoheadrightarrow_{\phi} \subseteq {}_n{\leftrightarrow}^{\circledast}_{\mu} \cdot \twoheadrightarrow^{=}_{\phi(\mu=n)} \cdot {}_{n'}{\leftrightarrow}^{\circledast}_{\mu'}$ with $\mu \geq n$ (and $\mu + \mu' \geq n' + n$), where $\phi(\mathsf{true})$ denotes $\curlyvee\phi$, the set of steps $\prec$-related to ϕ, and $\phi(\mathsf{false})$ denotes the set of all steps.

▶ **Proposition 36.** *If $\to$ is a random descent strategy for $\twoheadrightarrow$, then $\to$ is (ordered) compatible with $\twoheadrightarrow$ iff it is (ordered) decreasingly compatible with $\twoheadrightarrow$.*

Proof. For the only–if-direction, set $\prec$ to the empty relation and μ, n both to $\bot$. For the if-direction, let $\to$ have random descent and be (ordered) decreasingly compatible with $\twoheadrightarrow$. We show for $\to$-terminating a, that if $a \twoheadrightarrow_{\phi} b$ then $a \; {}_k{\leftrightarrow}^{\circledast}_{\lambda} \; b$ (with $\lambda \geq k$), by well-founded induction on the pair $(\mathsf{d}(a), \phi)$ ordered by the lexicographic product of $>$ and $\succ$, using 'vertical pasting' (cf. Fig. 5). We load the induction hypothesis to show that λ is finite. ◀

We present a simple, easily applicable sufficient criterion. (Note that it is weaker than Theorem 34 as it requires more, than needed, for objects *not* convertible to normal form.)

▶ **Corollary 37.** *If $\to$ is a random descent strategy with respect to the length measure, for $\twoheadrightarrow$ and $\prec$ a well-founded order $\prec$ on the steps of $\twoheadrightarrow$ such that for all $\twoheadrightarrow$-steps ϕ, it holds $\twoheadrightarrow_{\phi} \subseteq (\to \cdot \twoheadrightarrow^{=}_{\curlyvee\phi}) \cup (\to \cdot \twoheadrightarrow^{=} \cdot \leftarrow)$, then $\to$ is hyper-normalising and $\twoheadrightarrow$ has NF.*

Proof. By Theorem 34 and Proposition 36, instantiating measures to 0/1. ◀

▶ **Example 38** (‡)**.** To prove that the spine strategy $\to$ is hyper-normalising for $\beta(\eta)$ in the λ-calculus, it suffices to show the assumptions of Corollary 37 are satisfied when setting $\twoheadrightarrow$ to (nonempty) $\beta(\eta)$-multisteps, taking as well-founded order $\prec$ on them the *development* order, generated by ordering a multistep above each of its residuals after contracting anyone of its redex-patterns. This is a well-founded order by the finite developments theorem (see [2]). Let $\phi : M \mathrel{-\!\!\circ\!\!\twoheadrightarrow} N$ be a nonempty $\beta(\eta)$-multistep and $\psi : M \to M'$ a spine step.

If the respective (sets of) redex-patterns of ϕ and ψ do not have overlap, then we may compute the residual of each after the other with common reduct, say, N'. As spine redex-patterns are (uniquely) preserved after taking residuals, we have $\psi/\phi : N \to N'$. Observing that the 'other' residual ϕ/ψ either is empty (covered by reflexivity) or a nonempty multistep, we conclude to the 'right case' of Corollary 37.

If the redex-pattern of ψ has overlap with some redex-pattern, say ϕ_1, in ϕ, then we may develop ϕ as $M \to_{\phi_1} M' \twoheadrightarrow_{\phi/\phi_1} N$. We conclude to the 'left case' of Corollary 37 as ϕ/ϕ_1 is either empty or $\prec$-smaller than ϕ.

6 Hyper-normalisation for left–outer Dyck systems

We turn the reasoning in Example 38 into a critical pair criterion for a general class of term rewrite systems comprising both combinatory logic and the $\lambda\beta$-calculus. We dub the criterion *left–outer Dyck*, as it will entail (by Theorem 19) that all critical peaks can be completed into Dyck diagrams by means of *critically left–outer* steps, i.e. by left–outer steps that are closed under substitutions and (left–outer) contexts. The latter have RD so the criterion guarantees (by Theorem 34) hyper-normalisation for the left–outer strategy, cf. [21, Section 9].

We formalise our results in the setting of higher-order term rewrite systems where terms are simply typed λ-terms modulo $\alpha\beta\eta$-equality over simply typed (variables and) function symbols [23], using η-long β-normal forms as unique representatives of $\alpha\beta\eta$-equivalence classes of terms. To obtain decidability of criticality of left–outer steps, we focus on Nipkow's higher-order pattern rewrite systems [10], restricted to systems that are left-normal and local.

▶ **Definition 39.** A term is a pattern [12] if the free variables in it have sequences of pairwise distinct bound variables as arguments. A pattern is

- *linear* if each free variable that occurs in it, occurs in it exactly once;
- *fully-extended* [6] if each free variable occurring in it has as arguments a sequence comprising the variables bound above it;
- *local* [15, Footnote 1] if it is both linear and fully-extended;
- *left-normal* [14, 9] if each free position in it only $\prec_{lo}$-relates (see Definition 27) to other such, with positions in subterms having a free variable as head being *free*.

These notions extend to rewrite rules and systems via their left-hand side(s).

▶ **Example 40.** The higher-order rewrite rules corresponding to β- and η-reduction are:

$$@(\lambda x.M(x))N \to M(N) \qquad \lambda x.@Mx \to M$$

where λ and @ (henceforth left implicit) are appropriately typed function symbols, usually called abs respectively app [10, 19]. The left-hand sides $(\lambda x.M(x))N, \lambda x.Mx$ are patterns: the free variables M and N only have sequences of pairwise distinct bound variables, x and the empty sequence (twice), as arguments. The former, $(\lambda x.M(x))N$, is both local and left-normal: linear as its free variables M, N both occur once in it, fully-extended since M has the variable x bound above it as argument and N has no arguments, and left-normal as the free position $1 \cdot 2 \cdot 2 \cdot 1$ (of $M(x)$) only $\prec_{lo}$-relates to 2 (the position of N). The latter, $\lambda x.Mx$, is neither local nor left-normal. It is linear but not fully-extended since M does not have the variable x bound above it among its (empty) list of arguments, and not left-normal as the free position $2 \cdot 1 \cdot 1 \cdot 2$ (of M) $\prec_{lo}$-relates to $2 \cdot 1 \cdot 2$ (the position of x, not free). That is, the β-rule on its own constitutes a left-normal and local higher-order pattern rewrite system, but not so (locality falters) combined with the η-rule.

On first-order terms locality coincides with linearity. Throughout we use $\rightarrow$ and $\multimap\!\!\rightarrow$ to denote the one-step respectively (nonempty) multistep [19] abstract rewrite system underlying a local and left-normal higher-order pattern rewrite system. We focus on *left–outer* rewriting, denoted by $\to$, i.e. the restriction of $\rightarrow$ to contracting redexes at $\prec_{lo}$-more positions:

▶ **Definition 41.** For $\prec$ a strict order on positions, a position q of a redex-pattern is *$\prec$-more* if there is overlap between the redex-pattern and each redex-pattern at a position $p \prec q$, and *$\prec$-most* if there is no redex-pattern at a position $p \prec q$.

That $\to$ indeed is a strategy for $\rightarrow$ holds since any term not in normal form contains a *leftmost–outermost* redex, i.e. a redex at $\prec_{lo}$-most position. A left–outer redex need not be leftmost–outermost, e.g. when overlapped from above by a redex that *is* leftmost–outermost. Locality allows to characterise (critically) left–outer steps via (critically) left–outer contexts.

▶ **Definition 42.**

- A context is *left–outer* if it is single-hole and there is no redex-pattern at a position that $\prec_{lo}$-relates to the position of the hole, cf. [21, Definition 17];
- A context C is *critically* left–outer if for every substitution σ and left–outer context D, the context $D[C^\sigma]$ is left–outer. A step in such a context is a *critically left-outer* step.

Any critically left–outer context, in particular the empty context $\Box$, is left–outer. The context $f(x,\Box)$ is left–outer, but not critically so if $f(a,y) \rightarrow \ldots$, since instantiating x by a turns it into a non-left–outer context. Similarly, $f(\Box)$ is not critically left–outer if $g(f(x)) \rightarrow \ldots$.

▶ **Proposition 43.**

- *q is left–outer in $C[\ell^\sigma]_q$, for ℓ a left-hand side, iff C is left–outer;*
- *$q \cdot p$ is left–outer in $D[(C[\ell^\sigma]_p)^\tau]_q$, if D is left–outer and C is critically left–outer;*
- *A context is critically left–outer iff it is a single-hole context such that each symbol at a position that $\prec_{lo}$-relates to the position of the hole, is not a free variable and cannot be overlapped with a redex-pattern.*

Proof.

- This follows from that if a term contains a redex at position p, then changing it, e.g. replacing a subterm by a hole or vice versa, at any position $p \prec_{lo} q$ not overlapping the redex-pattern, does not change redexhood, by locality.
- Using substitutions are homomorphic, $D[(C[\ell^\sigma]_p)^\tau]_q = D[C^\tau[(\ell^\sigma)^\tau]_p]_q = D[C^\tau[\ell^{\sigma;\tau}]_p]_q$. By the assumption that D is left–outer and C is critically left–outer, $D[C^\tau]$ is left–outer. Combining both, we conclude by the previous item.
- By locality and left-normality. ◀

In the first-order case the proposition yields a decision procedure for whether or not a rewrite step is critically left–outer, since testing for the presence of variables in terms and unification of (parts of) left-hand sides of rules with terms/contexts are both effective. In the higher-order case this is not immediate in general: although (parts of) left-hand sides of rules are assumed to be patterns, the term/context may be arbitrary, a non-pattern. We use $\to_\#$ to denote the restriction of the left–outer strategy $\to$ to critically left–outer steps.

▶ **Definition 44.** A rewrite system $\rightarrow$ is *left–outer Dyck* if each critical peak[7] can be completed into a Dyck diagram by a conversion $\leftrightarrow^{\circledast}$ comprising $\to_\#$-steps only, for a given rule measure. Here a *rule* measure is a measure only depending on the rule applied.

[7] We employ a symmetric notion of critical *peak* arising from the, usually asymmetrically defined, notion of critical *pair*, by allowing either step of the peak to be the/a root step of the pair.

The $\lambda\beta$-calculus and Combinatory Logic are left–outer Dyck, in the absence of critical pairs. An ARS is locally Dyck iff its associated TRS is left–outer Dyck (measures *are* rule measures). Closure under contexts and substitutions makes rule measures suited for critical pair criteria.

▶ **Lemma 45.** *If $\twoheadrightarrow$ is left–outer Dyck, then $\to$ is locally Dyck.*

Proof. Consider a local peak of left–outer steps $t \leftarrow \cdot \to s$. By totality of $\leq_{lo}$, the positions of the respective contracted redexes are either identical or one is $<_{lo}$-related to the other. In either case the redex-patterns must have overlap, giving rise to a critical peak $t' \leftarrow \cdot \to s'$ of steps that are again left–outer such that the peak is encompassed via, say, left–outer context C and substitution σ, i.e. such that $t = C[t'^{\sigma}] \leftarrow \cdot \to C[s'^{\sigma}] = s$. By the assumption that the system is left–outer Dyck, for the given rule measure, the critical peak can be completed into a Dyck diagram by a $\to_{\#}$-conversion from t' to s'. By definition and by rule measures only depending on the rule applied, encompassing the conversion into the (left–outer) context C and substitution σ, yields a $\to$-Dyck-conversion from $t = C[t'^{\sigma}]$ to $C[s'^{\sigma}] = s$, as desired. ◀

▶ **Corollary 46.** *If $\twoheadrightarrow$ is left–outer Dyck, then $\to$ has random descent.*

Proof. By Lemmata 24 and 45. ◀

▶ **Example 47.** The left-normal, local term rewrite system with rules

$$a \twoheadrightarrow b \qquad f(x) \twoheadrightarrow g(x) \qquad h(f(b)) \twoheadrightarrow c \qquad c \twoheadrightarrow d \qquad h(g(b)) \twoheadrightarrow d$$

is left–outer Dyck, hence the left–outer strategy has random descent. The only interesting critical peaks are between the second and third rules, $h(g(b)) \twoheadleftarrow h(f(b)) \twoheadrightarrow c$. The peak is non-trivial but shown to be *root balanced joinable* [21] by $h(g(b)) \to_{\#} d \leftarrow_{\#} c$.

▶ **Example 48.** The left-normal, local term rewrite system with rules

$$a \twoheadrightarrow g(a) \qquad f(a) \twoheadrightarrow f(c) \qquad g(x) \twoheadrightarrow d \qquad c \twoheadrightarrow d$$

is left–outer Dyck, hence the left–outer strategy has random descent. Only overlap between the first and the second rule (and the other way around) gives rise to an interesting critical peak $f(g(a)) \twoheadleftarrow f(a) \twoheadrightarrow f(c)$. The peak is completable into a Dyck diagram but is *not* root balanced joinable: the redex-patterns contracted in the joining valley $f(g(a)) \to_{\#} f(d) \leftarrow_{\#} f(c)$ occur in the critically left–outer but non-empty, context $f(\Box)$.

Left-normality and locality can be viewed as syntactic conditions guaranteeing that left–outer redexes are *external* [8], they descend [19] uniquely until overlapped by the contracted redex:

▶ **Proposition 49.** *Left–outer steps descend along non-overlapping multisteps.*

Proof. Locality guarantees that being a redex or not only depends on its pattern, not on its variables being instantiated appropriately. This prevents creating a redex above the left–outer redex by steps below or parallel to its redex-pattern, just by changing instantiation of its variables. Left-normality guarantees that no redex-pattern can be created above a left–outer redex by means of contracting a redex parallel to (to the right of) it. ◀

▶ Remark. Non-fully-extendness of the η-rule causes that a left–outer redex u may descend to a non-left–outer redex along a non-left–outer step in $\lambda\beta\eta$, e.g. in $\lambda x.u((\lambda y.z)x)x \twoheadrightarrow \lambda x.uzx$.

▶ **Theorem 50.** *If a rewrite system is left–outer Dyck, then it has the normal form property, and the left–outer strategy is hyper-normalising for it.*

Proof. By assumption and Corollary 46 the left–outer strategy $\to$ has random descent. To conclude that the rewrite system $\rightarrowtriangle$ has the normal form property and $\to$ is hyper-normalising for it, it suffices to show that the same hold for the the (nonempty) multistep rewrite system $\mathrel{\circ\!\!\!\rightarrowtriangle}$ because $\rightarrowtriangle \subseteq \mathrel{\circ\!\!\!\rightarrowtriangle} \subseteq \rightarrowtriangle^{*}$. To that end, it suffices by Theorem 34 to show that $\to$ is ordered compatible with the rewrite system $\rightarrowtriangle$ and compatible with its converse $\leftarrowtriangle$, which in turn follow by Proposition 36 from that $\to$ is ordered decreasingly compatible with $\mathrel{\circ\!\!\!\rightarrowtriangle}$, and decreasingly compatible with its converse, which we both show simultaneously by well-foundedly ordering multisteps by the development order (see Example 38), considering an arbitrary nonempty multistep $W : t \mathrel{\circ\!\!\!\rightarrowtriangle} s$. By $\to$ being a strategy there is a left–outer step from t, say $u : t \to_m t'$. We distinguish cases on whether or not u overlaps some step in W.

If u overlaps no step in W, then both are orthogonal and we may compute their mutual residuals $W/u : t' \mathrel{\circ\!\!\!\rightarrowtriangle} s'$ and $u/W : s \mathrel{\circ\!\!\!\rightarrowtriangle} s'$. By Proposition 49, left–outer redex-patterns are (uniquely) preserved after taking residuals, so $u/W : s \to_m s'$ by rule measures only depending on the rule. Thence $\mathrel{\circ\!\!\!\rightarrowtriangle}_W \subseteq \to_m \cdot \mathrel{\circ\!\!\!\rightarrowtriangle}^{=} \cdot {}_m\!\leftarrow$, from which the conditions for (ordered) decreasing compatibility of $\to$ with (the converse of) $\mathrel{\leftarrowtriangle\!\!\!\circ}_W$ follow.

If the redex-pattern of u has overlap with some redex-pattern, say w, in W, then we may develop W as $w : t \rightarrowtriangle t''$ followed by $W' : t'' \mathrel{\circ\!\!\!\rightarrowtriangle} s$ with $W' = W/w$. Since the redex-patterns u, w yielding the peak $t' \leftarrow \cdot \rightarrowtriangle t''$ have overlap in its source t, the peak encompasses some critical peak $r' \leftarrow \cdot \rightarrowtriangle r''$, say via context C and substitution σ. The context C is left–outer as a prefix of the left–outer context in which u occurs. By the assumption that $\rightarrowtriangle$ is left–outer Dyck, the peak and its symmetric version can be completed into Dyck diagrams by $\to_{\#}$-conversions from r' to r'' and vice versa, respectively. Encompassing these again by the left–outer context C and substitution σ yields, by the steps in the conversions being *critically* left–outer and by rule measures only depending on the rule applied, Dyck diagrams for $\to$-conversions from t' to t'' and vice versa. By the former, $\mathrel{\circ\!\!\!\rightarrowtriangle}_W \subseteq \to_m \cdot {}_n\!\leftrightarrow_{\mu}^{\circledast} \cdot \mathrel{\circ\!\!\!\rightarrowtriangle}$ with $m + \mu > n$ from which we conclude to ordered decreasing compatibility of $\to$ with $\mathrel{\circ\!\!\!\rightarrowtriangle}_W$. By the latter, $\mathrel{\leftarrowtriangle\!\!\!\circ}_W \subseteq \mathrel{\leftarrowtriangle\!\!\!\circ}_{W'} \cdot \leftrightarrow^{\circledast}$ from which we conclude to decreasing compatibility of $\to$ with $\mathrel{\leftarrowtriangle\!\!\!\circ}_W$, as W is larger than W' in the development order. ◀

▶ **Example 51.** The rewrite systems of Examples 47 and 48, $\lambda\beta$- and CL-reduction are left-normal, local and locally Dyck, so the left–outer strategy is hyper-normalising for each.

▶ **Example 52.** Consider the local, left-normal first-order term rewrite system given by rules

$$\mathsf{zeros} \rightarrowtriangle_1 0 : \mathsf{zeros} \qquad \mathsf{hd}(x : y) \rightarrowtriangle_1 x \qquad \mathsf{hd}(\mathsf{zeros}) \rightarrowtriangle_2 0$$

with measures as indicated. Its critical peak $0 \;{}_2\!\leftarrowtriangle \mathsf{hd}(\mathsf{zeros}) \rightarrowtriangle_1 \mathsf{hd}(0 : \mathsf{zeros})$ is completed by $0 \;{}_1\!\leftarrowtriangle \mathsf{hd}(0 : \mathsf{zeros})$ into a Dyck diagram, so the left–outer strategy is hyper-normalising.

7 Conclusion

We have generalised (hyper-)normalisation results from [20, 21] using the random descent property from [16]. Our development is based on a clean separation between the abstract and term rewrite results. At the abstract level we have introduced novel methods to compare strategies, by Dyck diagrams, and to prove their (hyper-)normalisation, by compatibility. At the term level, we have introduced a class of higher-order term rewrite systems, left–outer Dyck systems, comprising $\lambda\beta$ and CL.

Theorem 50 generalises [7, Theorem 25] on which their further developments are based. The generalisation is proper in that our results are restricted neither to deterministic strategies nor to first-order term rewrite systems, while sharing the advantage of being completely local (even more so). We expect it to be possible to incorporate several techniques from their work, in particular basic normalisation and factorisation,[8] into our approach. The approach to normalisation due to [11] is mostly incomporable to ours. On the one hand, that approach is based on square permutations (a special case of random descent) and on having finite 'permutation equivalence' classes (not needed in our approach). On the other hand, there normalisation for notions of result other than normal forms (think of head-normal forms) are considered. We intend to apply approach to hyper-normalisation to, e.g., λ-calculi with explicit substitutions or the necessary strategy [18], and compare our results to those of [5].

Acknowledgements. We thank B. Felgenhauer, J. Nagele, the attendants of the TCS seminar (Amsterdam) and the Master Seminar (Innsbruck), and the FSCD reviewers for feedback.

References

1. F. Baader and T. Nipkow. *Term Rewriting and All That*. Cambridge University Press, 1998.
2. H.P. Barendregt. *The Lambda Calculus: Its Syntax and Semantics*, volume 103 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, 2nd revised edition, 1984.
3. H.P. Barendregt, J.R.K. Kennaway, J.W. Klop, and M.R. Sleep. Needed reduction and spine strategies for the lambda calculus. *Information & Computation*, 75(3):191–231, 1987. `doi:10.1016/0890-5401(87)90001-0`.
4. N. Dershowitz. On lazy commutation. In *Languages: From Formal to Natural*, volume 5533 of *Lecture Notes in Computer Science*, pages 59–82. Springer, 2009. `doi:10.1007/978-3-642-01748-3_5`.
5. J.R.W. Glauert, R. Kennaway, and Z. Khasidashvili. Stable results and relative normalization. *Journal of Logic and Computation*, 10(3):323–348, 2000. `doi:10.1093/logcom/10.3.323`.
6. M. Hanus and C. Prehofer. Higher-order narrowing with definitional trees. In *Proceedings of the 7th International Conference on Rewriting Techniques and Applications*, volume 1103 of *Lecture Notes in Computer Science*, pages 138–152. Springer, 1996. `doi:10.1007/3-540-61464-8_48`.
7. N. Hirokawa, A. Middeldorp, and G. Moser. Leftmost outermost revisited. In *Proceedings of the 26th International Conference on Rewriting Techniques and Applications*, volume 36 of *Leibniz International Proceedings in Informatics*, pages 209–222, 2015. `doi:10.4230/LIPIcs.RTA.2015.209`.
8. Gérard Huet and Jean-Jacques Lévy. Computations in orthogonal rewriting systems, I. In *Computational Logic: Essays in Honor of Alan Robinson*. The MIT Press, 1991. (accessed 27-4-2016). URL: `http://pauillac.inria.fr/~levy/pubs/81robinson1.pdf`.
9. J.W. Klop. *Combinatory Reduction Systems*. PhD thesis, Rijksuniversiteit Utrecht, 1980. (accessed 27-4-2016). URL: `http://janwillemklop.nl/Jan_Willem_Klop/Bibliography_files/9.PhDthesis-total.pdf`.
10. R. Mayr and T. Nipkow. Higher-order rewrite systems and their confluence. *Theoretical Computer Science*, 192(1):3–29, 1998. `doi:10.1016/S0304-3975(97)00143-6`.

[8] But note that for a rewrite system $\twoheadrightarrow$ given by $a \twoheadrightarrow b \twoheadrightarrow c$, $a' \twoheadrightarrow b' \twoheadrightarrow c'$, $a \twoheadrightarrow a'$, $b \twoheadrightarrow b'$ and $c \twoheadrightarrow c'$, and a strategy $\to$ obtained by omitting the step from a to b, factorisation fails, for $a \twoheadrightarrow b \to c$, but our methods, in particular Theorem 50, do apply.

11 P.-A. Melliès. A stability theorem in rewriting theory. In *Proceedings of the 13th Annual IEEE Symposium on Logic in Computer Science*, pages 287–298. IEEE Computer Society Press, 1998. doi:10.1109/LICS.1998.705665.

12 D. Miller. Unification of simply typed lambda-terms as logic programming. In *Proceedings of the 8th International Conference on Logic Programming*, pages 253–281. The MIT Press, 1991. (accessed 27-4-2016). URL: http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/iclp91.pdf.

13 M. Newman. On theories with a combinatorial definition of "equivalence". *Annals of Mathematics*, 43(2):223–243, 1942. doi:10.2307/1968867.

14 M.J. O'Donnell. *Computing in Systems Described by Equations*, volume 58 of *Lecture Notes in Computer Science*. Springer, 1977. doi:10.1007/3-540-08531-9.

15 V. van Oostrom. Finite family developments. In *Proceedings of the 7th International Conference on Rewriting Techniques and Applications*, volume 1232 of *Lecture Notes in Computer Science*, pages 308–322. Springer, 1997. doi:10.1007/3-540-62950-5_80.

16 V. van Oostrom. Random descent. In *Proceedings of the 18th International Conference on Rewriting Techniques and Applications*, volume 4533 of *Lecture Notes in Computer Science*, pages 314–328. Springer, 2007. doi:10.1007/978-3-540-73449-9_24.

17 V. van Oostrom. Confluence by decreasing diagrams – converted. In *Proceedings of the 19th International Conference on Rewriting Techniques and Applications*, volume 5117 of *Lecture Notes in Computer Science*, pages 306–320. Springer, 2008. doi:10.1007/978-3-540-70590-1_21.

18 R.C. Sekar and I.V. Ramakrishnan. Programming in equational logic: Beyond strong sequentiality. In *Proceedings of the 5th Annual IEEE Symposium on Logic in Computer Science*, pages 230–241. IEEE Computer Society Press, 1990. doi:10.1109/LICS.1990.113749.

19 Terese. *Term Rewriting Systems*, volume 55 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 2003.

20 Y. Toyama. Strong sequentiality of left-linear overlapping term rewriting systems. In *Proceedings of the 7th Annual IEEE Symposium on Logic in Computer Science*, pages 274–284. IEEE Computer Society Press, 1992. doi:10.1109/LICS.1992.185540.

21 Y. Toyama. Reduction strategies for left–linear term rewriting systems. In *Processes, Terms and Cycles: Steps on the Road to Infinity: Essays Dedicated to Jan Willem Klop on the Occasion of His 60th Birthday*, volume 3838 of *Lecture Notes in Computer Science*, pages 198–223. Springer, 2005. doi:10.1007/11601548_13.

22 F. Winkler and B. Buchberger. A criterion for eliminating unnecessary reductions in the Knuth–Bendix algorithm. In *Proceedings of the Colloquium on Algebra, Combinatorics and Logic in Computer Science, Volume II*, volume 42 of *Colloquia Mathematica Societatis J. Bolyai*, pages 849–869, 1986.

23 D.A. Wolfram. *The Clausal Theory of Types*, volume 21 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1993.

24 H. Zantema, B. König, and H.J.S. Bruggink. Termination of cycle rewriting. In *Proceedings of the 25th International Conference on Rewriting Techniques and Applicationsand the 12th International Conference on Typed Lambda Calculi and Applications*, Lecture Notes in Computer Science, pages 476–490. Springer, 2014. doi:10.1007/978-3-319-08918-8_33.