Reaction to the Wikipedia page on Newman's Lemma

Vincent van Oostrom
University of Sussex
vvo@sussex.ac.uk

March 4, 2026

Abstract

This is a reaction to the current Wikipedia page on Newman's Lemma,[1] in particular to the characterisation due to Eriksson on it. I make suggestions for improving the page, recognisable by the word Wikipedia in blue and by the use of first-person singular. Though the latter deviates from standard scientific practice, it is based on suggestions on work by others being personal. The results on which the suggestions are based, supporting the narrative, are presented in the standard scientific style though, as they are known from the literature.

Newman's Lemma $\neq$ Newman's Lemma. On Wikipedia Newman's Lemma is stated to be a result for *abstract rewriting systems*, where the Wikipedia page on the latter defines such a system to be a set equipped with an endorelation.

My first reaction is that Newman's paper [13] presenting the result known as Newman's Lemma does *not* deal with relations or orders. Newman consciously steered clear from those as is clear from his [13, footnote 3]. Instead, he made use of systems having steps as first-class-citizens, with every step having a source object and a target object. He stated the result discussed here [13, Theorem 3] in terms of such systems, not in terms of relations.

For that very reason, we have taken (starting from [23], see Definition 8.2.2) to using *rewrite system* to refer to the systems introduced by Newman, and to using *rewrite relation* only for the particular rewrite systems where there is at most 1 step from a to b, for any a and b.

Using that terminology, the title of this paragraph turns from a contradiction into a truth: the result proved by Newman was a result for rewrite *systems*, whereas at many places in the literature, including Huet's paper and Wikipedia, Newman's Lemma is stated as a result for rewrite *relations*. These are not the same simply because they concern different structures (systems *vs.* relations). Thus, where it is stated on [6, p. 800] of Huet's paper that the result there ... *appears in its full generality in* ... Newman's paper, it should have stated that it appeared there in *more* generality, since Newman proved his result for rewrite systems. I suggest to add an explanation on Wikipedia expressing the distinction between rewrite relations and rewrite systems and to make clear that the result proved by Newman was for the latter, so different from what is usually referred to as Newman's Lemma.

The difference / relationship between rewrite systems (as introduced by Newman using topological notions) and rewrite relations (as used in the 1980s by Huet and others) is the same as that between categories and quasi-orders. Indeed, the latters can be (formally) seen as the formers equipped with extra structure (units on objects and composition of steps, satisfying the monoid laws [21]). For the same reason that categories and quasi-orders are not conflated in the literature one shouldn't conflate rewrite systems[2] and rewrite relations, in rewrite theory.[3] In my opinion, the confusion that was caused in the rewriting literature by conflating both, has been detrimental to the field, if only by causing a Babylonian confusion of tongues, and should not be perpetuated.

---

[1]`https://en.wikipedia.org/wiki/Newman's_lemma`, accessed 4th of March 2026 see the appendix.

[2]Employed in rewriting by authors like Newman, Hindley, Plotkin, Stark, Melliès, Khasidashvili, ...

[3]Employed in rewriting by authors like Dershowitz, Jouanaud, Huet, Klop, ...

Remark 1. Unfortunately, I also contributed to the confusion initially, before finding that conflating rewrite systems and rewrite relations was at the root of the problems I was encountering, and settling on the above distinction and corresponding nomenclature, keeping the notions apart ever since in my writings.

Remark 2. Of course, by rewrite systems generalising rewrite relations, many results transfer between them, but that should in my opinion then be made explicit. By (local) confluence and termination transferring between a rewrite system $\to$ and its associated rewrite relation $\to'$ (obtained by forgetting the identity of steps), Newman's Lemma for a rewrite system $\to$ is equivalent to the same result for its associated rewrite relation. However, results for rewrite systems depending on the identity of steps, *e.g,* on strategies and on steps overlapping, typically refine those for rewrite relations, if they can be stated at all for the latter. For instance, the classical notion of *critical pair* for rewrite relations is refined by the notion of *critical peak* for rewrite systems, and in Combinatory Logic there is no way to even state whether or not $I\,(I\,a) \to I\,a$ adheres to the outermost strategy at the level of the rewrite relation associated to its term rewrite rule $\iota\,x : I\,x \to x$, but in its associated rewrite system [23, Chapters 8 and 9] it can be stated and is meaningful since there there are *two steps* (both witnessing that $I\,(I\,a)$ is rewritable to $I\,a$): the step $\iota\,(I\,a)$ contracts the outermost I-redex-pattern so is outermost, but the step $I\,(\iota\,a)$ is not as it contracts the innermost I-redex-pattern.

Newman's Lemma $\neq$ Diamond Lemma. On Wikipedia it is stated that Newman's Lemma is a lemma for abstract rewriting systems and *commonly* called the **Diamond Lemma**.

Not so in the textbooks on rewriting [2, 1, 14, 23]. These do define[4] the *diamond property* to be $\leftarrow \cdot \to \,\subseteq\, \to \cdot \leftarrow$, *i.e.* that for every $a \leftarrow \cdot \to b$ there is a $a \to \cdot \leftarrow b$ constituting a diagram that is a *diamond*. But these books *nowhere* refer to Newman's Lemma as the Diamond Lemma. Hence I suggest to change wording on Wikipedia to reflect that.

Newman's Lemma provides just one set of sufficient (but not necessary) conditions (termination and local confluence of $\to$) for reductions $\twoheadrightarrow$ to have the diamond property, among several other sets of conditions that are at one's disposal for abstract rewrite systems, like *random descent* [13] and *decreasing diagrams* [15] (which is a *complete* [16, 7] method for showing $\twoheadrightarrow$ to have the diamond property), both discussed below. That makes referring to Newman's Lemma as *the* Diamond Lemma also confusing in my opinion.

Remark 3. The focus of rewrite theory being on single steps, its nomenclature typically refines that of other fields. Local confluence and its diagram are therefore also known as the *weak* Church–Rosser property (WCR) respectively a *weak* diamond [2].

Huet did not give the first account of a simple proof of Newman's Lemma. The proof of Newman's Lemma ([13, Theorem 3]) given on the Wikipedia page is that of Lemma 24 in Huet's seminal 1980 paper [6]. In [6], Huet discusses and dismisses several other proofs of Newman's Lemma (including the one of Newman) known at the time. However, in 1976, 4 years prior to his paper, a simple proof of Newman's Lemma had been given already in the famous[5] *Blue Preprint* [3]. We refactor that simple proof (of [3, Lemma II.1.11]), using current notations and terminology [23], obtaining Newman's Lemma as a corollary.

Definition 1. A *fork in the road*[6] is a peak $b \twoheadleftarrow a \twoheadrightarrow c$ where b, c are not joinable; a is its *root*.

Observe that neither reduction in a fork in the road can be empty.

Lemma 1. *If a is the root of some fork in the road for a locally confluent rewrite system $\to$, then $a \to a'$ for some a' that is the root of a fork in the road.*

---

[4]See [2, Definition 3.1.11], [1, Definition 2.7.8], [14, Definition 2.4.1], [23, Definition 1.1.8(v)].

[5]That preprint can be seen as being a stepping stone for Barendregt's 1980 standard book [2] on the λ-calculus, where the proof appears as Proposition 3.1.25.
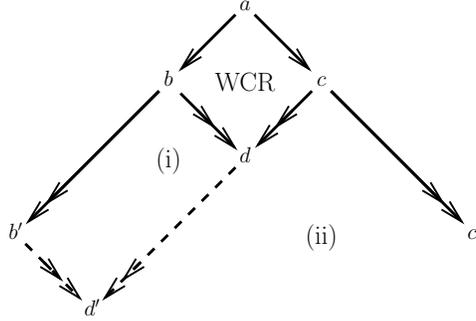
[6]This is *ad hoc* terminology.

Figure 1: Proof by picture of Lem. 1

Proof. By assumption and the observation there is a fork in the road $b' \twoheadleftarrow b \leftarrow a \rightarrow c \twoheadrightarrow c'$ for some b, b', c, c' with b' and c' non-joinable. By local confluence there is a valley $b \twoheadrightarrow d \twoheadleftarrow c$. We distinguish cases on whether or not $b' \twoheadleftarrow b \twoheadrightarrow d$ is a fork in the road, see Fig. 1. (i) If it is, we may set $a' := b$. (ii) Otherwise, b' and d are joinable so there is a valley $b' \twoheadrightarrow d' \twoheadleftarrow d$ for some d'. But then $d' \twoheadleftarrow d \twoheadleftarrow c \twoheadrightarrow c'$ must be a fork in the road, as otherwise b' (via its reduct d') and c' would be joinable, so we may set $a' := c$. $\qquad\square$

This proof / the proof in the Blue Preprint, in [3], is not constructive; though either b or c must be a fork in the road, it may not be decidable which one is. Huet's proof can be seen as a constructive version of it for terminating systems.

Corollary 1 (Newman's Lemma). *If $\rightarrow$ is locally confluent and terminating, then it is confluent.*

Proof. If $b \twoheadleftarrow a \twoheadrightarrow c$ were not joinable, a fork in the road, then by Lem. 1 there would be an infinite reduction from a through roots of forks in the road, contradicting termination. $\qquad\square$

Wikipedia states that *a number of other proofs exist* (other than the proof presented there due to Huet). To more accurately reflect history, I suggest to (also) state on Wikipedia that the Blue Preprint had a *simple proof* of Newman's Lemma at that time, or to simply present that proof / the above lemma.

Remark 4. I learned rewriting theory through the work of my PhD supervisor, Jan Willem Klop, If anything is the legacy of his work, it is that approaching rewrite theory from the perspective of *infinite* reductions often provides *more insight*. That pertains to his famous Δ-example, to his studies of standardisation, infinitary rewriting, streams *etc.*, but also to the above proof of Newman's Lemma. It is based on a *coinductive* construction rather than an *inductive* one and has the added bonus of characterising forks in the road, that from them infinite reductions may be constructed *only going through other such forks*, just from assuming local confluence. Lem. 1 pertains also to rewrite systems that are non-terminating. It yields a contradiction only when assuming termination, as in Newman's Lemma.

Its analysis can be and has been refined further. For instance, [23, Lemma 14.1.8] shows that when tiling a finite peak with local confluence diagrams does not terminate, there must be an infinite reduction through infinitely many so-called *splitting points*; *cf.* also, *e.g.*, [9],[26, Example 11].

More generally, in rewriting the notion of a potentially infinite reduction is the main derived notion, with finite reductions arising as such having a target. In my opinion, there's often an undue focus on finite reductions in the literature, with some literature not even considering the notion of an infinite reduction (only reductions having a target are considered). This may be caused by a focus on the classical notion of *computation* over the more recent notion of *interaction*, with a major difference between both being that typically computations *do* terminate whereas

interactions *do not*, *e.g.*, an operating system does not terminate, its goal is not to compute a value but to run forever, to keep on interacting.

Wikipedia works, but slowly. Relatively recently a paragraph was added to the Wikipedia page on Newman's Lemma, on a result due to Eriksson [4, 5] (see below). Despite having worked in the area for 30 years and knowing the result (and of strengthenings of it; see below), I was until now not aware of Eriksson's work in rewriting and of the literature cited by him, so I would like to express my gratitude to the contributor who included it on the Wikipedia page. *Vice versa*, that contributor (and also Eriksson), do (did) not seem to know about closely related results in the (rewriting) literature, that had and have been developed before and after. If I were to guess the reason for these developments having proceeded independently until now, I would say that this is due to the respective areas of game theory and rewriting sharing only few practitioners. (I for one, know next to nothing about combinatorial game theory.) In any case, I suggest to update Wikipedia to reflect also the rewriting literature on this, discussed below, not just Eriksson's notions and characterisation (and also not just its special case for termination systems).

Eriksson's strong convergence $\iff$ random descent. In Corollary 2.1 of [13], Newman gave a sufficient (see below) local condition for there being 'random descent' [13, § 5] in a rewrite system. The idea for his description apparently was that if that's the case, if there is random descent, then *descending* (rewriting) *randomly* (following any strategy) is guaranteed to reach the normal form, if there is one, and at minimal cost (always by the same number of steps) at that. Eriksson in [4, p. 5][5, p. 380] dubbed that very same property *strong convergence* (SC), apparently unaware of Newman's work and (weaker) result on it.[7] Around the same time as Eriksson, also Toyama established a result [24] on random descent, weaker than Eriksson's, but stronger than Newman's, see Fig. 4, but without citing Newman (for that) and without giving the property a name.[8] To honour Newman, we named, in turn ignorant of Eriksson's work, an equivalent property *random descent* in [17, 22] inspired by [13, 24].

Definition 2. A rewrite system $\to$ has *random descent* (RD) [22, Definition 22] if for every b, if there is a conversion of difference n from b to some normal form a, then there is a reduction from b to a of length n, where the *difference* [24][22, Example 6] of a conversion is its number of forward $\to$-steps minus its number of backward $\leftarrow$-steps. If we only require $\leq n$ in that, we speak of the *ordered normal form* (ONF) property [22, Definition 12].

Note the difference of a reduction is just its length. ONF trivially entails that if an object is
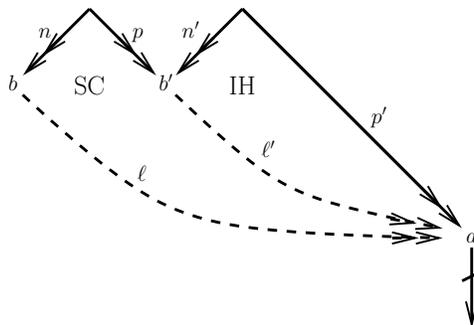

Figure 2: Proof by picture of SC $\implies$ RD; $\ell =_{\text{SC}} (-n) + p + \ell' =_{\text{IH}} (-n) + p + (-n') + p'$

---

[7]Though Eriksson does cite Newman's paper [13] in his PhD thesis [4, p. 6], he only does so to refer to Newman's Lemma, *i.e.* to [13, Theorem 3], not to Newman's notion of random descent appearing just before it in that same section [13, § 5].

[8]That Eriksson and Toyama do not refer to each other is less of a surprise, as they developed their ideas roughly contemporaneously.

terminating then all reductions to normal form are of the same length and end in the same normal form [22, Lemma 24], *i.e.* Eriksson's SC, by considering the peak between any two maximal reductions from the object and asymmetry of $\leq$. Conversely, SC entails RD (hence ONF by reflexivity of $\leq$) as can shown by induction on the number of peaks in a conversion and illustrated in Fig. 2, analogous to how the Church–Rosser property is shown equivalent to confluence [13].

Below, we will stick to RD and ONF and will mostly avoid SC. More generally, since the terminology (*strong*) *convergence* was already imported [23] into rewriting in its classical topological sense.[9], I suggest Wikipedia to steer clear from it, on pages discussing rewriting, such as the page on Newman's Lemma, or at least make clear that it's terminology from a different field, and not used in rewriting for the reason mentioned. I also suggest to include references to the work on RD, as that notion is more easily applicable than SC.

Eriksson's polygon property is not isolated. The account on Wikipedia of Eriksson's polygon property suggests it is isolated from properties in rewrite theory. It is not. To illustrate that, we recapitulate various conditions on local peaks known from the rewriting literature, both before and after [4, 5], and relate them to the polygon property.

Definition 3. A rewrite system $\to$ such that for every local peak $b \ _\phi\!\leftarrow \cdot \to_\psi c$, for steps ϕ, ψ,

- $\phi = \psi$, is *deterministic* [23, Definition 9.1.4];

- there is a valley $b \to \cdot \leftarrow c$, has the *diamond* property (see above);

- either $b = c$ or there is a valley $b \to \cdot \leftarrow c$ [13, Theorem 1], has the *sub*-diamond property;

- there is a valley $b \to^n \cdot \ ^n\!\leftarrow c$ for some natural number n, is *balanced* weakly Church–Rosser (bWCR) [24, Definition 3.1];

- either b and c are not terminating or there is a valley $b \to^n \cdot \ ^n\!\leftarrow c$ for some natural number n, has the *polygon* property [4, p. 5][5, p. 380];

- either b is not terminating or there is a valley $b \to^n \cdot \ ^m\!\leftarrow c$ for some natural numbers $n \geq m$, is *ordered* weakly Church–Rosser (OWCR) [17, p. 324] or *ordered* locally confluent [22];

- there is a conversion from b to some d that is either not terminating or equal to c with every prefix of the conversion having a non-negative difference, is *locally Dyck* [22, Definition 16].

Remark 5. The list of properties on local peaks in Def. 3 comprises the most important ones, but is not exhaustive. Some related ones are mentioned on [4, p. 6] (the *Jordan–Dedekind* property requiring valleys to have reductions of length *2*, trivially entailing bWCR), and at the bottom of [17, p. 324] (all trivially entailing OWCR but not necessarily the polygon property).

Remark 6. Newman used but didn't name the condition we cleped *sub-diamond* above in his [13]. Sub-diamond amounts to what is called *strong* confluence in [10, Proposition 1], but we avoid that terminology since it is already in use [1, 23] in the more general sense (asking only $b \twoheadrightarrow \cdot \leftarrow^= c$) introduced by Huet before [6, p. 801]. Previously, I called it *linear orthogonal* in [8, Definition 4], but in hindsight I think also that naming is inappropriate since a linear orthogonal rewrite system need not be *orthogonal* in the sense of [23, § 8.7], it need not have the so-called *cube* property; it need not entail that every local 3-peak $a \to b_i$ for $i \in \{1, 2, 3\}$ can be completed into a cube by 6 diamonds, for the same reason that the diamond property need not [18, p. 22].

Except for deterministism entailing only sub-diamond, not diamond, each property trivially entails the subsequent one, see Fig. 3. For instance, that the polygon property entails ordered local confluence (OWCR) follows by taking n and m equal, and that ordered local confluence entails local Dyck follows by setting d to b if that is not terminating, and to c otherwise (which works since any prefix of a valley $b \to^n \cdot \ ^m\!\leftarrow c$ with $n \geq m$ clearly has a non-negative difference).

---

[9]Even overloaded more in some rewriting literature [1] by letting it denote completeness, *i.e.* termination and confluence.
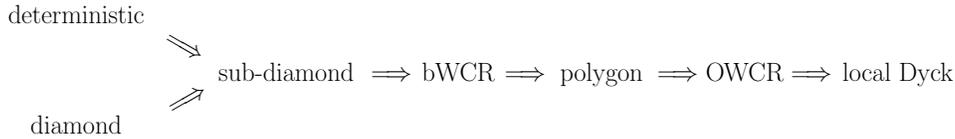
Figure 3: Trivial implications for properties of local peaks in Def. 3
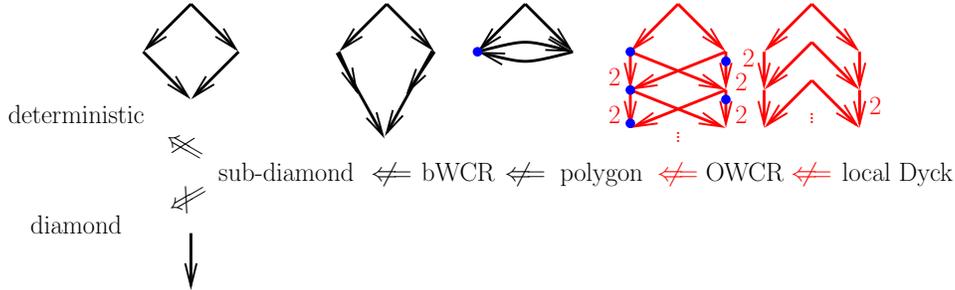

Figure 4: (Local) failure of implications for properties of local peaks in Def. 3

The converses of the implications in Fig. 3 fail as illustrated by the rewrite systems in Fig. 4, where to the right of the polygon the failures are only failures of locality, indicated in red, by which we mean that the property to the right of the witnessing rewrite system can be verified to hold *locally* via reductions / conversions of *bounded* size (in the examples size 8 suffices), whereas verifying the property to its left cannot (in principle requires *unbounded* exploration to establish non-termination). For instance, that OWCR does not entail the polygon property locally, and that the polygon property does not entail bWCR, can both be seen for these examples using a simple parity argument: objects that are reachable from the target of the left step of the local peak at the top in an *even* number of steps (indicated by the blue dots • in the figure) can only be reached from the target of its right step in an *odd* number of steps (and *mutatis mutandis* the same *vice versa*), so never by the *same* number of steps, never by reductions of the *same* length.

As for the *global* property of random descent, the literature on the *local* properties in Def. 3 is largely disjoint. On the one hand, Eriksson's work on the polygon property in the 90s and later papers based on it in combinatorial game theory till today do not cite the rewriting literature on the other properties. On the other hand, the rewriting literature from the 1940s till today on the other properties in Def. 3, does not cite the combinatorial game literature on the polygon property. More precisely, Eriksson missed in [4, 5] the properties *above* the polygon property in Def. 3 / to the *left* of it in Fig. 3, and I / we missed in [25, 17, 22] the polygon property.

Wikipedia not only misses all rewriting literature on the local properties in Def. 3 other than the polygon property, from 1942 till today, but incorrectly defines Eriksson's polygon property to be what is known as Toyama's balanced weak Church–Rosser property, with the former but not the latter allowing for that both b and c are non-terminating in the local peak in Def. 3. Therefore, I suggest Wikipedia to include a discussion of the above properties other than the polygon property and to refer to [24] not to [4, 5] and name the property bWCR (in the way the result is stated, at present).

Remark 7. Since for its application to terminating rewrite systems given on Wikipedia, bWCR and the polygon property are *equivalent* (trivially so, because of the absence of non-terminating objects), attribution to both Toyama [24] or Eriksson [4] would be appropriate for this instance.[10]

---

[10]I don't know who discovered their result first, though by being contemporaneous, it stands to reason that their discoveries were independent.

Local Dyck characterises random descent. Wikipedia states without proof, that for terminating systems Eriksson established: polygon property $\implies$ SC. Apart from that Eriksson showed a more general statement (not just bWCR $\implies$ SC; see above), he in fact gave a *characterisation* [4, p. 5,6][5, p. 380,381] of the *global* property SC by means of the *local* polygon property, *i.e.* he showed SC $\iff$ polygon property without assuming termination. I suggest to state his result on Wikipedia and also present the proof, as it can be presented easily pictorially in the style of the pictorial proof of Newman's Lemma there, *cf.* also the pictorial proofs of [24, Lemma 3.2], [17, Theorem 2], [22, Theorems 18 and 19].

Wikipedia misses that progress has been made since Eriksson's characterisation, by both *weakening* his *local* assumption, the polygon property, to the system being local Dyck, and *strengthening* his *global* conclusion, SC, to random descent. More precisely, instantiating [22, Lemma 24] for the special case of the length measure, yields: local Dyck $\iff$ random descent. I suggest to update Wikipedia with at least that statement, and the definition of the notions. Whether it would be appropriate to also present the proof of the statement I don't know. The proof of [22, Lemma 24] is rather roundabout so not directly suitable; it factors through the proof of a more general *commutation* (see below) result presented there [22, Lemma 14,Theorem 19], and moreover works for arbitrary measures not just the length measure. To allow a reader to make an informed decision about that (suitability), we give a direct proof below obtained by performing 'cut-elimination' on the instantiation of commutation to confluence and the instantiation to the length measure. Though the resulting proof is already simpler, it still is not quite straightforward.

Lemma 2. *Local Dyck $\iff$ random descent.*

Proof. We claim the following special case of ONF holds: for every b, for every reduction of length n from b to some normal form a, every Dyck conversion from b can be extended by a *reduction* to a into a conversion of difference $\leq n$, where a conversion is *Dyck* if every prefix of it has a non-negative difference. Since the claim is a generalisation of SC[11] we conclude to random descent.

We prove the claim by induction on n first and the length of the conversion second. If the conversion from b is empty, we conclude trivially by extending it by the reduction from b to a. Otherwise, it has shape $b \leftrightarrow^* c \leftrightarrow c'$ for some c, c'. By the induction hypothesis for the reduction $b \twoheadrightarrow a$ (for the same n) and the (shorter) prefix $b \leftrightarrow^* c$, say of difference k, there is a reduction $c \twoheadrightarrow a$ of length $\leq n - k$, which is non-negative by being a reduction and smaller than or equal to n since the conversion being Dyck, the difference k of its prefix is non-negative. We distinguish cases on $c \leftrightarrow c'$:

- If $c \leftarrow c'$, then extending the conversion $b \leftrightarrow^* c \leftarrow c'$ by the reduction $c' \rightarrow c \twoheadrightarrow a$ is as desired, with difference $\leq n$ as the consecutive steps $c \leftarrow c' \rightarrow c$ cancel out in it.

- If $c \rightarrow c'$, then the reduction $c \twoheadrightarrow a$ is of shape $c \rightarrow b' \twoheadrightarrow a$ for some b'. By local Dyck for the local peak $b' \leftarrow c \rightarrow c'$, there is a Dyck conversion $b' \leftrightarrow^* d$, say of difference k', for some d that is either not terminating or equal to c'. We split cases on the disjunction, in both cases using that the induction hypothesis applies to the reduction $b' \twoheadrightarrow a$ of length $\ell \leq (n - k) - 1$ and an arbitrary Dyck conversion from b' since the length of the former is $\ell < n - k \leq n$:

  That d is non-terminating cannot hold, since then we could construct a conversion $b' \leftrightarrow^* d \twoheadrightarrow d'$ of arbitrary difference ℓ', in particular of difference $\ell' > \ell$. This would lead to a contradiction since by the induction hypothesis for (the reduction $b' \twoheadrightarrow a$ and) that conversion then would entail the existence of a reduction $d' \twoheadrightarrow a$ of *negative* length $\leq \ell - \ell'$.

  If $d = c'$, then the induction hypothesis for the reduction $b' \twoheadrightarrow a$ (of length ℓ) and the conversion $b' \leftrightarrow^* c'$ (of difference k') leads to extending the latter into a conversion of difference $\leq \ell$ by a reduction $c' \twoheadrightarrow a$ (of length $\leq \ell - k'$)l see Fig. 5. To see that the difference of the extended conversion $b \leftrightarrow^* c \rightarrow c' \twoheadrightarrow a$ is bounded as desired, we compute $n \geq k + (n - k) \geq k + 1 + \ell \geq k + 1 + k' - k' + \ell \geq k + 1 + (\ell - k')$. $\qquad\square$

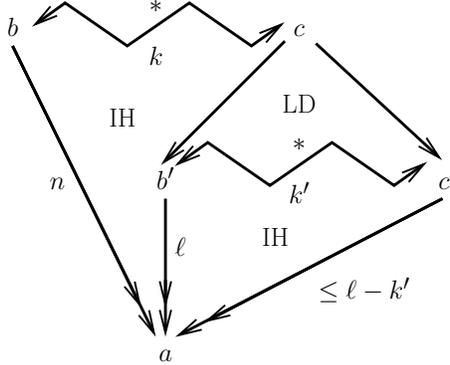[11] And of the *ordered peak normal form* (OPN) property, equivalent to RD for the length measure [22, Lemma 24].

Figure 5: Proof by picture of final case in the proof of Lem. 2

Since the polygon property $\implies$ local Dyck $\implies$ random descent $\implies$ SC as seen above, Eriksson's result [4, 5] that the polygon property $\iff$ SC, entails that all of them are equivalent. This could give the impression that nothing was gained by Lem. 2. However, that would be mistaken: as seen above, local Dyck does not locally entail the polygon property as illustrated in Fig. 4, SC only globally entails random descent (induction on the conversion was needed for that; see Fig. 2).

Example 1. *Consider the rewrite system $\to$ having steps $a \to b$, $a \to c$, and $b \to c$, $c \to b$ as displayed in the middle in Fig. 4. As critical peaks it has $b \leftarrow a \to c$ and its converse.*

OWCR / local Dyck is shown to hold locally *by the (degenerate) valley $b \leftarrow c$ and mutatis mutandis the same for the converse critical peak, entailing random descent. Observe that (very)* local *search suffices to find these valleys; in fact, single steps suffice. However, the lengths of the reductions from b respectively c to a common reduct are* never *the same, as they always have distinct parity as argued above. As a consequence, for establishing the polygon property for that local peak directly one would need some kind of* global *reasoning, establishing that b / c are non-terminating. In the case of this example that is still quite trivial, can* also *be done locally, but the rewrite systems to its right in Fig. 4 illustrate that in general non-termination might be hard to establish, may need proper* global *reasoning.*

Remark 8. An interesting side-effect of the above is that it opens up not only establishing random descent locally, by local Dyck as in the example, but also checking *non-termination* locally, without having to come up with an infinite reduction, only by checking how critical peaks can be completed: If *some* local peak was completed into a diagram by means of a Dyck conversion of difference different from 0,[12] then *all* of the objects on the diagram are non-normalising. This pertains to the critical peak in the previous example.[13][14]

The correct setting? Reflecting on his characterisation, Eriksson writes in his PhD thesis [4, p. 6]: *The possibility of infinite paths and infinite graphs, while we demand equal length of all finite maximal paths, is natural and is, in our opinion, the correct setting.* The question is of course, the correct setting for what?

For the purpose of comparing strategies and establishing random descent, I suggest that *ordered local commutation* (OLCOM) and its instance *ordered local confluence* (OWCR) provide the better

---

[12]Something that can't happen when checking the polygon property, but may checking OWCR / local Dyck.

[13]Of course, this is not a complete method for determining non-termination; *e.g.*, deterministic non-terminating systems need not have such peaks.

[14]There are other methods to establish non-termination based on *peaks*. For instance, in a rewrite system having the diamond property, any object that is the source or target of a step is non-terminating. In other words, only isolated objects are then terminating; that is, only normal forms are.

setting. The *local* property OLCOM was introduced in [17] to guarantee and characterise (in certain cases) the *global* property that one strategy is *better* (reaches the normal form in fewer steps) than another [17, Theorem 2]. It was developed to solve the open problem whether the F_∞ strategy is a maximal and perpetual strategy for a certain λ-calculus with explicit substitutions, in the affirmative [17, Example 7]. OLCOM distinguishes itself from the polygon property in that it is *asymmetric* in requiring the length of the left leg of a diagram to be *greater than or equal to* that of its right leg, whereas the polygon property is *symmetric* in demanding (see Eriksson's reflection above) both legs to have *equal* length. Obviously, such asymmetric commutation diagrams provide the correct setting for *comparing* strategies,[15] *not* the symmetric confluence diagrams considered by Eriksson.[16] However, we think that, contrary to Eriksson's claim, not the symmetric polygon property but the asymmetric ordered local confluence / OWCR property, obtained by instantiating both systems in the asymmetric OLCOM commutation diagrams by the same rewrite system, provides the correct setting also for the characterisation of random descent / Eriksson's SC. This view is supported by the results in [17, 22] recounted above, showing that OWCR has better local applicability than the polygon property (as was illustrated in Fig. 4 for the case of confluence).[17] Intuitively, this holds since OWCR refines the analysis of the symmetric notion of equality $=$ (between legs of the valleys for the local peak) in the polygon property, via the asymmetric notion of inequality $\geq$ in OWCR (between the legs of the *two* valleys for the local peak and its converse), with the conjunction of $\geq$ and its converse $\leq$ entailing $=$ by anti-symmetry.

In [22], we generalised the set-up of [17] in two ways. First, we generalised the (asymmetric) commutation diagrams of [17] by allowing for (certain) conversions instead of either the peak or valley in the diagrams, leading to the characterisation chain of [22, Figure 3], reproduced here in Fig. 6, introducing both the ordered Church–Rosser property and the local Dyck property, characterising the former global property by the latter local property. This makes the setting
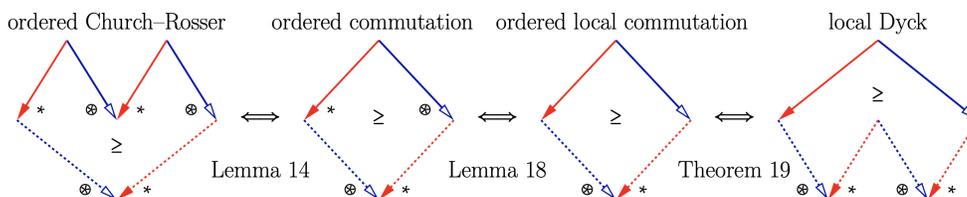

Figure 6: Characterisation chain of the ordered Church–Rosser property [22, Figure 3]

already better than the one of [17] by local Dyck having better local applicability than OLCOM (as was illustrated in Fig. 4 for the case of OWCR / confluence) and by the ordered Church–Rosser property being more global than ordered commutation. Second, we generalised comparing reductions by their *length* into comparing them via their *weight*, allowing to assign weights to steps[18] and weighing a reduction by the sum (in some ordered monoid) of the weights of its steps. This makes the set-up better again as it allows for the characterisation [19, Theorem 1][20, Theorem 3]: uniformly complete $\iff$ random descent, for *some assignment of weights*, where a rewrite system is *uniformly* complete [19, Definition 1] if all objects convertible to normal form are complete, *i.e.* confluent and terminating. This has the following characterisation of completeness of as immediate corollary [19, Corollary 2][20, Corollary 1]:

Corollary 2. *SN & CR $\iff$ WN & OWCR. That is, a rewrite system is terminating and confluent iff it is normalising and ordered locally confluent (for some assignment of weights).*

---

[15]Recently, this has been confirmed in [11], where it was shown how to apply (in our terminology, but developed independently by them) ordered commutation to show correctness of program transformations and optimisations, much along the lines of our ISR course [12] based on [17, 22].

[16]Conversely, we expect OLCOM to be applicable in combinatorial game theory.

[17]Though we didn't aim for this when developing [17]; it is a fortunate side-effect.

[18]Note that for this it is essential to have steps at one's disposal, *i.e.* to work with rewrite systems; rewrite relations are insufficient.

Remark 9. This fails in the unweighted case (when weighing all steps by 1), as illustrated by the terminating and confluent rewrite system $c \leftarrow a \rightarrow b \rightarrow c$; unweighted OWCR fails trivially.

[20, §III.A] illustrates the power of weights and of these characterisations by making short shrift of various results both novel and classical. For instance, Church and Rosser's classical [23, Theorem 4.8.5] result that in the λI-calculus any term that is normalising is terminating follows from showing random descent to hold for an appropriate assignment of weights [20, Example 14], and so does Klop's classical [23, Theorem 1.2.3(iii)] result: WCR & WN & Inc $\implies$ SN [20, Example 13].

Remark 10. As far as I know, Cor. 2 provides the first alternative to Newman's Lemma for establishing completeness by analysing local peaks. As for the unweighted case above, also the weighted version of OWCR / local Dyck yields a local method to show *non-termination* of an object: if it is on a local diagram having a non-trivial weight, then it is non-terminating.

Combining the innovations of [22] with the above ones from [17], suggests that the most correct setting is that of *weighted asymmetric conversion diagrams (for two rewrite systems)*, though it is not clear to me yet how to present it in a palatable way on Wikipedia.

Remark 11. A technical innovation of [22] over [17] is the representation of *infinite* reductions as *steps* in a derived rewrite system [22, Definition 10]. The purpose of that was to express properties like OLCOM and OWCR by means of a *single* diagram, instead of by cases for finite and infinite reductions separately. Though it technically works out fine, the jury is still out on whether that provides the correct setting.

On the same page? Wikipedia presents Eriksson's characterisation (or rather one direction of it restricted to terminating systems; see above) on the same page as Newman's Lemma. I suggest (with the above proviso) to have a separate page for the narrative on random descent, including Eriksson's characterisation, since both that characterisation and Newman's Lemma are important but incomparable: one assumes order constraints on diagrams but not termination whereas the other assumes termination but no order constraints; neither entails the other.

Remark 12. It might be natural to include Cor. 2 on the Wikipedia page of Newman's Lemma since both are characterisations of completeness, of being terminating and confluenct, by criteria on local peaks, with SN and WCR in Newman's Lemma being respectively stronger than WN and weaker than OWCR. However, this then would suggest that the page be on completeness rather than on Newman's Lemma. Though that might be appropriate conceptually, it still might not be a good idea as it might make Newman's Lemma lose prominence.

Wikipedia pages on characterisation chains? To have a *chain* of equivalences as in Fig. 6, characterising the ordered Church–Rosser property by means of the local Dyck property (via ordered commutation and ordered local commutation), is a recurring pattern[19] in rewriting theory: the characterisation[20] of the Church–Rosser property by means of local decreasing diagrams with respect to conversions (via confluence and local decreasingness) in [26] adheres to the same pattern, as does the characterisation [23, Exercise 1.3.12] due to Winkler and Buchberger of SN & CR by SN & locally confluent *below the source* (via SN & confluence and SN & WCR, *i.e.* Newman's Lemma), *cf.* [26], and there may be others. I suggest not only that each of these chains is worthy of having its own Wikipedia page, but also that it might be worthwhile to have a Wikipedia page discussing the pattern (visualised in Fig. 6) each of the chains adheres to, discussing their commonality.

---

[19]Formally described in the introduction of [22].

[20]Whether this is a complete characterisation [16, Conjecture 2.3.31] is still open, but for all practical (countable) cases it is [16, Corollary 2.3.30], a result that was recently and impressively strengthened by Ivanov [7].

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, 1984.

[3] H.P. Barendregt, J. Bergstra, J.W. Klop, and H. Volken. Degrees, reductions and representability in the lambda calculus. Preprint 22, Utrecht University, Department of Mathematics, 1976. (The Blue Preprint). URL: `https://dspace.library.uu.nl/handle/1874/15119`.

[4] K. Eriksson. *Strongly Convergent Games and Coxeter Groups*. PhD thesis, Kungl Tekniska Högskolan, Stockholm, 1993. URL: `https://archive.org/details/eriksson-strongly-convergent-games-thesis`.

[5] K. Eriksson. Strong convergence and a game of numbers. *European Journal of Combinatorics*, 17(4):379–390, 1996. `doi:10.1006/eujc.1996.0031`.

[6] G. Huet. Confluent reductions: Abstract properties and applications to term rewriting systems. *Journal of the ACM*, 27(4):797–821, 1980. `doi:10.1145/322217.322230`.

[7] I. Ivanov. Completeness of the decreasing diagrams method for proving confluence of rewriting systems of the least uncountable cardinality. In *FSCD*, volume 337 of *LIPIcs*, pages 25:1–25:20, 2025. `doi:10.4230/LIPIcs.FSCD.2025.25`.

[8] Z. Khasidashvili, M. Ogawa, and V. van Oostrom. Uniform normalisation beyond orthogonality. In *RTA*, volume 2051 of *LNCS*, pages 122–136, 2001. `doi:10.1007/3-540-45127-7_11`.

[9] J.W. Klop, V. van Oostrom, and R. de Vrijer. A geometric proof of confluence by decreasing diagrams. *Journal of Logic and Computation*, 10(3):437–460, 2000.

[10] Y. Lafont. Interaction nets. In *POPL*, pages 95–108, 1990. `doi:10.1145/96709.96718`.

[11] K. Muroya and M. Hamana. Term evaluation systems with refinements: First-order, second-order, and contextual improvement. In *FLOPS*, volume 14659 of *LNCS*, pages 31–61, 2024. `doi:10.1007/978-981-97-2300-3_3`.

[12] J. Nagele and V. van Oostrom. Commutation. Course in the Advanced Track of the International School on Rewriting (ISR), 2017. URL: `https://hzantema.win.tue.nl/isrnagele.pdf`.

[13] M.H.A. Newman. On theories with a combinatorial definition of "equivalence". *Annals of Mathematics*, 43:223–243, 1942. `doi:10.2307/2269299`.

[14] E. Ohlebusch. *Advanced Topics in Term Rewriting*. Springer, 2002.

[15] V. van Oostrom. Confluence by decreasing diagrams. *Theoretical Computer Science*, 126(2):259–280, 1994.

[16] V. van Oostrom. *Confluence for Abstract and Higher-Order Rewriting*. PhD thesis, Vrije Universiteit, Amsterdam, 1994.

[17] V. van Oostrom. Random descent. In *RTA*, volume 4533 of *Lecture Notes in Computer Science*, pages 314–328. Springer, 2007. `doi:10.1007/978-3-540-73449-9_24`.

[18] V. van Oostrom. Residuation = skolemised confluence. In *IWC*, pages 20–25, 2023. URL: `http://cl-informatik.uibk.ac.at/iwc/2023/proceedings.pdf`.

[19] Vincent van Oostrom. Uniform completeness. In *IWC 2022*, pages 19–24, 2022. URL: `http://cl-informatik.uibk.ac.at/iwc/2022/proceedings.pdf`.

[20] Vincent van Oostrom. Confluence and orthogonality by residuation illustrated by the problem of the calissons, 2025. URL: `http://www.javakade.nl/research/pdf/candobyres.pdf`.

[21] Vincent van Oostrom. Redeeming Newman, orthogonality in rewriting; past, present and future in a 1-algebraic setting. In *IWC 2025*, pages 45–51, 2025. URL: `https://iwc2025.github.io/IWC2025_proceedings.pdf#page=50`.

[22] Vincent van Oostrom and Yoshihito Toyama. Normalisation by random descent. In *FSCD*, volume 52 of *LIPIcs*, pages 32:1–32:18, 2016. `doi:10.4230/LIPIcs.FSCD.2016.32`.

[23] Terese. *Term Rewriting Systems*. Cambridge University Press, 2003.

[24] Y. Toyama. Strong sequentiality of left-linear overlapping term rewriting systems. In *LICS*, pages 274–284, 1992. `doi:10.1109/LICS.1992.185540`.

[25] 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 *LNCS*, pages 198–223, 2005. `doi:10.1007/11601548_13`.

[26] Oostrom. V. van. Confluence by decreasing diagrams, converted. In *RTA*, volume 5117 of *NCS*, pages 306–320, 2008. `doi:10.1007/978-3-540-70590-1_21`.

Appendix.

Newman's lemma

In theoretical computer science, specifically in term rewriting, **Newman's lemma**, also commonly called the **diamond lemma**, is a criterion to prove that an abstract rewriting system is confluent. It states that local confluence is a sufficient condition for confluence, provided that the system is also terminating. This is useful since local confluence is usually easier to verify than confluence.

The lemma was originally proved by Max Newman in 1942.[1][2] A considerably simpler proof (given below) was proposed by Gérard Huet.[3] A number of other proofs exist.[4]

Statement and proof

The lemma is purely combinatorial and applies to any relation. Owing to the context where it is commonly applied, it is stated below in the terminology of abstract rewriting systems (this is simply a set whose elements are called terms, equipped with a relation $\rightarrow$ called reduction, and see the corresponding article for definitions of termination, confluence, local confluence and normal forms).

Newman's lemma:[5][6][7][8] If an abstract rewriting system is terminating and locally confluent, then it is confluent. As a corollary, every term has a *unique* normal form.

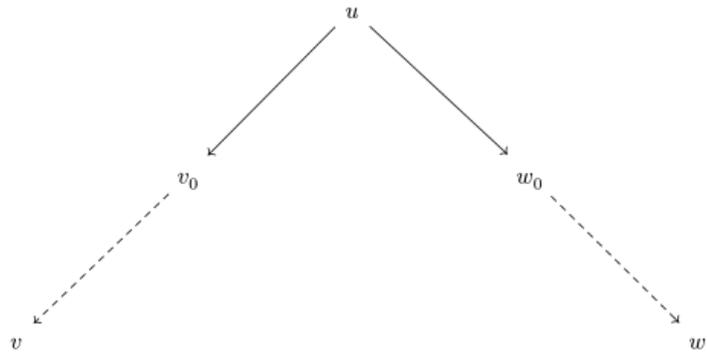**Proof:** We prove by well-founded induction on u along $\rightarrow$ that every diagram

$$
\begin{array}{ccc}
 & u & \\
\swarrow & & \searrow \\
v & & w
\end{array}
$$
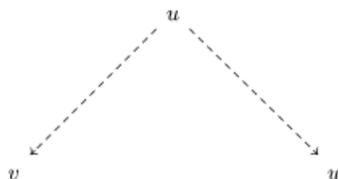
can be extended to a diagram

$$
\begin{array}{ccc}
 & u & \\
\swarrow & & \searrow \\
v & & w \\
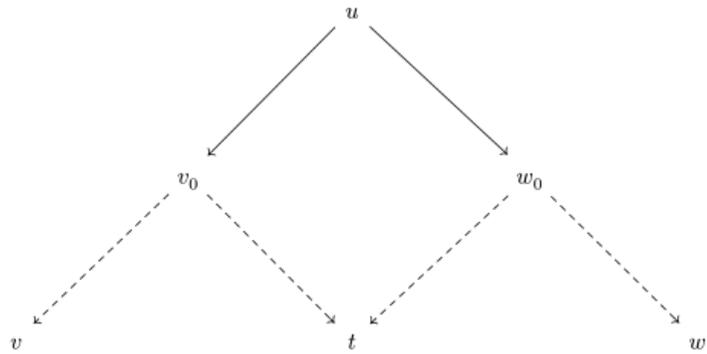\searrow & & \swarrow \\
 & t &
\end{array}
$$
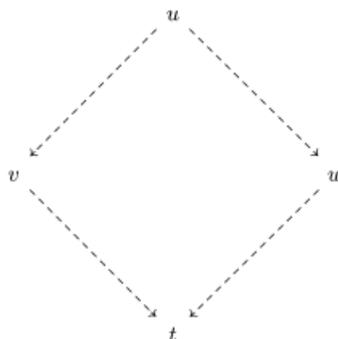
where the dotted arrows represent sequences of arbitrarily many reductions by $\rightarrow$.
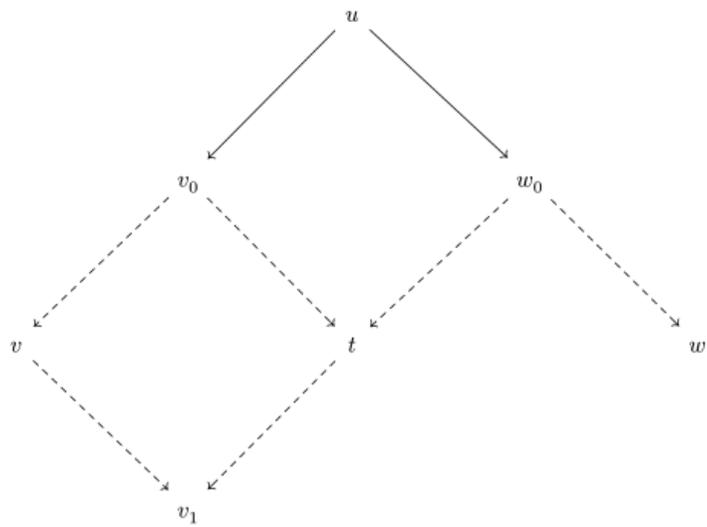
If $u = v$ or $u = w$, this is trivial. Otherwise, we have at least one reduction on each side:
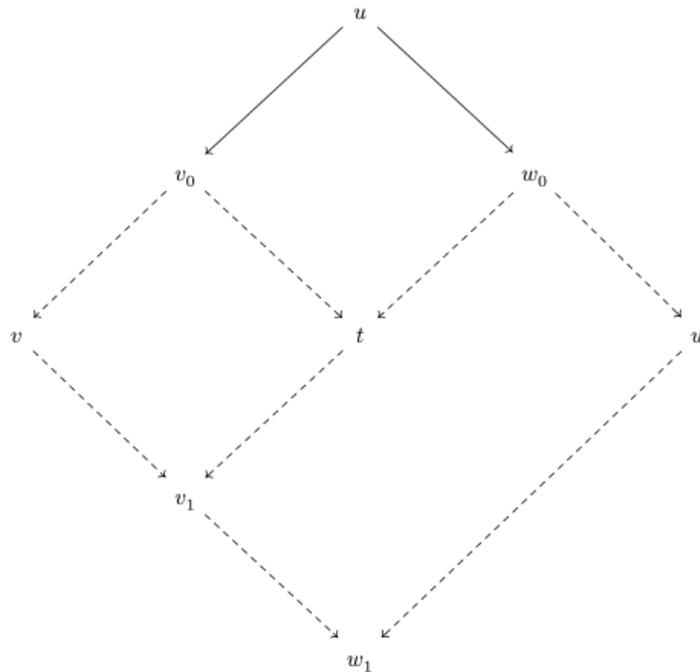
By local confluence, this diagram can be extended to:

then by induction hypothesis on v_0:

and finally, by induction hypothesis on w_0:

Eriksson's polygon property lemma

A related result was shown by Kimmo Eriksson in 1993.[9][10] Recall that an abstract rewriting system is locally confluent if for any two reductions $a \to b$ and $a \to c$, there exists d such that $b \to^* d$ and $c \to^* d$. If additionally it is required that the reduction chains $b \to^* d$ and $c \to^* d$ have the same length, then the system is said to have the **polygon property**. Examples of rewriting systems with the polygon property include bubble sort and the chip-firing game.

Eriksson's polygon property lemma shows that if an abstract rewriting system is terminating and has the polygon property, then not only is it confluent (according to Newman's lemma), but additionally every terminating chain of reductions from a given state has the same length.

References

1. Newman, Max (1942). "On theories with a combinatorial definition of "equivalence" ". *Annals of Mathematics*. **43** (2): 223–243.
2. van Oostrom, Vincent. "Newman's Proof of Newman's Lemma" (https://web.archive.org/web/20240415162426/http://cl-informatik.uibk.ac.at/users/vincent/research/publication/pdf/newmansproof.pdf) (PDF). Archived from the original (http://cl-informatik.uibk.ac.at/users/vincent/research/publication/pdf/newmansproof.pdf) (PDF) on April 15, 2024.
3. Huet, Gérard (1980). "Confluent reductions: Abstract properties and applications to term rewriting systems" (https://inria.hal.science/hal-04716458). *Journal of the ACM*. **27** (4): 797–821. doi:10.1145/322217.322230 (https://doi.org/10.1145%2F322217.322230).
4. Klop, Jan Willem (1990). "Term rewriting systems: From Church-Rosser to Knuth-Bendix and Beyond". *Automata, languages, and programming: 17th international colloquium*. Lecture Notes in Computer Science. Vol. 443. Warwick University, England: Springer. pp. 350–369. doi:10.1007/BFb0032044 (https://doi.org/10.1007%2FBFb0032044). ISBN 978-3-540-52826-5.

5. Baader, Franz; Nipkow, Tobias (1998). *Term Rewriting and All That*. Cambridge University Press. doi:10.1017/CBO9781139172752 (https://doi.org/10.1017%2FCBO9781139172752). ISBN 0-521-77920-0.

6. Terese (2003). *Term Rewriting Systems*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press.

7. Harrison, John (2009). *Handbook of Practical Logic and Automated Reasoning*. Cambridge University Press. p. 260. ISBN 978-0-521-89957-4.

8. Cohn, Paul Moritz (1980). *Universal Algebra*. D. Reidel Publishing. pp. 25–26. ISBN 90-277-1254-9.

9. Eriksson, Kimmo (1993). *Strongly Convergent Games and Coxeter Groups* (https://archive.org/details/eriksson-strongly-convergent-games-thesis) (Takn.dr thesis). Stockholm: KTH.

10. Eriksson, Kimmo (1996). "Strong Convergence and a Game of Numbers". *European Journal of Combinatorics*. **17** (4): 379–390. doi:10.1006/eujc.1996.0031 (https://doi.org/10.1006%2Feujc.1996.0031).

External links

- Diamond lemma (https://planetmath.org/diamondlemma) at PlanetMath.