A geometric proof of confluence by decreasing diagrams

Jan Willem Klop[1,2,3] Vincent van Oostrom[1,4] Roel de Vrijer[2]
jwk@cs.vu.nl oostrom@phil.uu.nl rdv@cs.vu.nl

[1] *CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands*
[2] *Department of Theoretical Computer Science, Vrije Universiteit de Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands*
[3] *Department of Computer Science, University of Nijmegen Toernooiveld 1, 6525 ED Nijmegen, The Netherlands*
[4] *Department of Philosophy, Utrecht University P.O. Box 80.126, 3508 TC Utrecht, The Netherlands*

ABSTRACT

The criterion for confluence using decreasing diagrams is a generalization of several well-known confluence criteria in abstract rewriting, such as the strong confluence lemma. We give a new proof of the decreasing diagram theorem based on a geometric study of infinite reduction diagrams, arising from unsuccessful attempts to obtain a confluent diagram by tiling with elementary diagrams.

2000 Mathematics Subject Classification: 68Q42, 52C20
1999 ACM Computing Classification System: F.1.1, F.4.2
Keywords and Phrases: abstract rewriting, tiling, confluence
Note: Work of first and second author carried out under project SEN2

1. Introduction

Abstract rewriting is the initial part of the theory of rewriting where objects have no structure and the rewrite relation is just a binary relation on the set of objects. Usually there is not just one rewrite relation, but an indexed family of rewrite relations present. There are several useful and well-known lemmas for such abstract rewrite systems that give conditions for confluence: Newman's Lemma [10], Huet's strong confluence lemma [7], Staples' request lemmas [14], the lemma of Hindley-Rosen [6].

A common generalization of all these lemmas has been obtained in van Oostrom [12, 11], elaborating an unpublished note of de Bruijn [5]. De Bruijn's original proof was a complicated nested induction, while van Oostrom used a certain invariant for the diagram construction called *decreasing diagrams*. A slightly different invariant called *trace-decreasing diagrams* was used in Bezem et al. [2]; this invariant will be used in the present paper. The theorem of de Bruijn and van Oostrom is concerned with labeled reductions. For a version of the theorem where points instead of edges are labeled, see Bognar [3], with a proof checked by the Coq proof checker.

In this paper we give a proof of this 'confluence by decreasing diagrams' theorem that is totally different from the two mentioned above. The proof is by an analysis of the geometry of, possibly infinite, reduction diagrams, resulting from two co-initial diverging finite reduction sequences, by 'tiling' with elementary reduction diagrams. Infinite diagrams arise this way when we have a failure of confluence.

Such infinite reduction diagrams are interesting geometric objects themselves; the simplest one is the diagram in Figure 1 that we will call the *Escher diagram*. In the sequel we will give several more examples of infinite reduction diagrams, some of them exhibiting an interesting fractal-like boundary, some of them reminiscent to the pictures of M.C. Escher, with a repetition of the same pattern, receding in infinity.

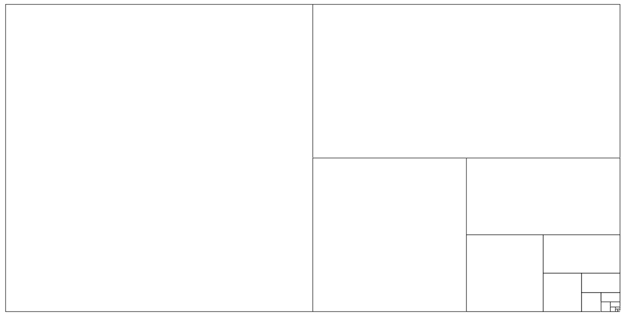

Figure 1: The Escher diagram

Actually, we consider an enrichment of mere reduction diagrams, namely diagrams with a 'tree covering'. A tree covering of a diagram determines an ancestor-descendant relation between the edges appearing in a reduction diagram. By means of a tree covering an edge can be traced back to its ancestor edge on one of the original divergent reduction sequences. The theorem proved in this paper states the impossibility of certain infinite diagrams with a tree-covering. Since an infinite reduction diagram composed of (trace-)decreasing diagrams would give rise in a natural way to a tree covering—of the impossible kind—we have as an immediate corollary then the theorem of confluence by decreasing diagrams. The method of proof of our theorem is purely geometric. It employs topological notions such as condensation points of point sets in the real plane.

2. Abstract reduction systems

An *abstract reduction system* (ARS) $\mathcal{A}$ is a set A equipped with a collection of rewrite or reduction relations $\to_\alpha$, indexed by some set I of indexes: $\mathcal{A} = \langle A, (\to_\alpha)_{\alpha \in I} \rangle$. The index set I is a well-founded partial order. In examples, we will use the set of natural numbers with the usual ordering as index set. The union of the rewrite relations $\to_\alpha$ is denoted by $\to$. We use the notation $\twoheadrightarrow$ for the transitive-reflexive closure of $\to$. Idem $\twoheadrightarrow_\alpha$ with respect to $\to_\alpha$.

The ARS $\mathcal{A}$ is called *confluent* or CR (Church-Rosser), if

$$\forall a, b, c \in A \; (a \twoheadrightarrow b \wedge a \twoheadrightarrow c \;\Rightarrow\; \exists d \in A \; (b \twoheadrightarrow d \wedge c \twoheadrightarrow d)).$$

A, non-equivalent, weaker version of CR is WCR: $\mathcal{A}$ is called *locally confluent* or WCR (weakly Church-Rosser), if

$$\forall a, b, c \in A \; (a \to b \wedge a \to c \;\Rightarrow\; \exists d \in A \; (b \twoheadrightarrow d \wedge c \twoheadrightarrow d)).$$

The CR and WCR properties are depicted respectively in Figure 2, the picture of WCR giving rise to the notion of *elementary diagram*, which will be defined in the next section.

It is well-known that in strongly normalizing ARSs (i.e., ARSs without infinite reduction sequences) we have WCR $\Rightarrow$ CR (Newman's Lemma [10]). The essence of Newman's lemma is that because of the strong normalization condition the process of tiling with elementary diagrams must terminate. The decreasing diagrams method studied in this paper amounts essentially to giving a weaker condition, yet yielding the termination of tiling.

For more on abstract reduction systems we refer to Klop [8].

3. Elementary diagrams

As said, the property WCR inspires the notion of elementary diagram, which we now define. It originates from Klop [9], where also the notion of *improper* elementary diagram was introduced.

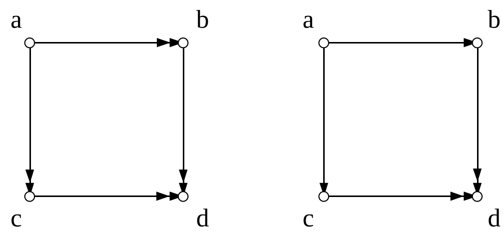

Figure 2: CR and WCR

Definition 1 An *elementary diagram* (e.d.) is a configuration of two reduction steps $a \to b$ and $a \to c$, issuing from the same object a, and reductions $b \to b_1 \to \cdots b_m \equiv d$ and $c \to c_1 \to \cdots c_n \equiv d$ that join b and c. (Note that m and/or n may be zero: if, e.g., $m = 0$ we have $b \equiv d$.)

This is the abstract notion of elementary diagram. An e.d. can be rendered geometrically as a rectangle with some nodes on its sides as in Figure 3 (from left to right we have in the first diagram $m = n = 1$, in the second $m = 3$, $n = 2$, in the third $m = 2$, $n = 0$ and in the last one $m = n = 0$).

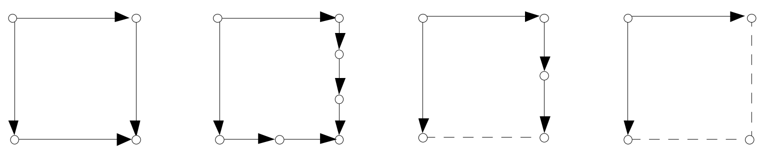

Figure 3: Elementary diagrams

To distinguish the abstract notion of e.d. in definition 1 from its geometric representation, we call the latter a *geometric e.d.* So in a geometric e.d. we have the original steps $a \to b$ and $a \to c$ as upper and left-hand sides, and the converging reductions $b \to b_1 \to \cdots b_m \equiv d$ on the right and $c \to c_1 \to \cdots c_n \equiv d$ on the lower side, d in the lower right corner. Objects are rendered as nodes, reduction steps as edges (with arrow). In case $m = 0$ or $n = 0$ the corresponding side is a so-called *empty step*, drawn as a dashed line. We need empty steps in order to keep our diagram constructions rectangular.

The geometric e.d.'s will be used as 'tiles' with the intention to obtain a completed *reduction diagram* as in Figure 6 (see the next section). To make this tiling process successful we need also e.d.'s with empty steps as upper or left-hand side. Hence empty steps give rise to trivial geometric e.d.'s as in Figure 4. Following Klop [9], we call these *improper e.d.'s*.

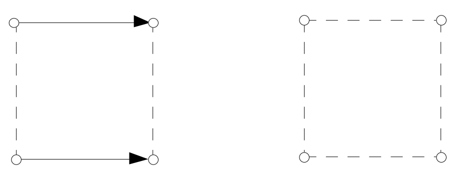

Figure 4: Improper elementary diagrams

We sum up the main characteristics of geometric e.d.'s and some conventions for dealing with them.

1. A geometric e.d. is a rectangle with some nodes on its sides.

2. Nodes represent objects. (These may be added as labels, but are mostly suppressed.)

3. Drawn *edges* connecting adjacent nodes represent reduction steps, always downward or from left to right. (This may or may not be indicated by an arrowhead and the index of the reduction step as label.)

4. The upper and left sides of a geometric e.d. are called its *initial sides*, the lower and right the *converging sides.*

5. There are nodes on the four corner points of the rectangle. The left-upper node is the *initial* node.

6. The converging sides may contain a finite number of extra nodes (the initial sides may not).

7. The converging sides may also consist of a dashed line (with no extra nodes on it); dashed lines represent empty steps, and therefore the nodes connected by a dashed line represent the same object.

8. In a geometric e.d. also one or both of the initial sides may be dashed; in that case the e.d. is called improper.

9. In an improper e.d. each converging side is identical to the opposite initial one.

10. Geometric e.d.'s are supposed to be *scalable*: they can be stretched or shrunk horizontally and vertically, as long as they keep their rectangular form. (Nodes on converging sides are in general placed equidistantly, but this is not essential.)

Now, since in this paper we look for sufficient conditions for the implication WCR $\Rightarrow$ CR, we will in fact assume WCR for all considered ARSs. This amounts to the following. Given an object a and reduction steps $a \rightarrow_\alpha b$ and $a \rightarrow_\beta c$, there is an e.d. with a as initial node and $a \rightarrow_\alpha b$ and $a \rightarrow_\beta c$ as initial steps. Geometrically this means that any configuration of a node with two adjacent edges, one to the left and one downward, but without converging sides—to be called an *open corner*—can be filled with an appropriate e.d. In the terminology of Bezem et al. [2], we have a *full set* of e.d.'s:

> The supply of tiles (geometric e.d.'s) is such for each open corner there is at least one fitting tile.

Note that because of the symmetry of WCR, with each tile $\mathcal{T}$ in our supply, we also have the tile that results by mirroring $\mathcal{T}$ on the diagonal through its initial node.

4. Reduction diagrams

4.1 Finite reduction diagrams

The e.d.'s are used as building blocks ('tiles') in the construction of reduction diagrams, in an attempt to construct for given (initial) finite reductions $a \twoheadrightarrow b$ and $a \twoheadrightarrow c$ issuing from the same object a, two convergent reductions $b \twoheadrightarrow d$ and $c \twoheadrightarrow d$. This attempt is successful if we arrive at a finite, *completed* reduction diagram as in the example of Figure 6.

The construction of such a diagram (by *tiling*) starts by setting out the initial reductions, one horizontally, one vertically, both starting at the same node representing a; and then proceeds by subsequently adjoining e.d.'s, as in Figure 5, at an *open corner* in which an e.d. can be fitted. (Here, by an *open corner* we mean a node with two adjacent edges, one to the left and one downward, without converging sides.) An e.d. fits if its initial node and initial sides match with the open corner, i.e., represent, respectively, the same object, and reduction steps with the same index from I. (The geometric fitting is accomplished by scaling the e.d.)

Each stage in the tiling process just described is called a *reduction diagram.* A reduction diagram is *completed* if it has no open corners. With a completed diagram the tiling process terminates. It is

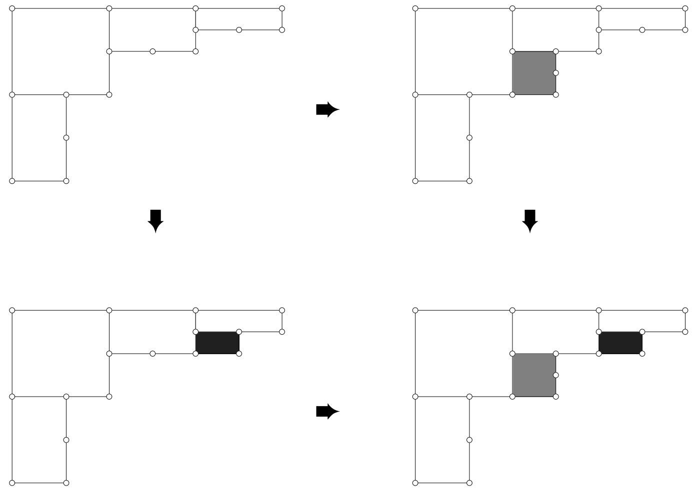

Figure 5: Adjoining e.d.'s to reduction diagrams

easy to see that a completed reduction diagram with initial reductions $a \twoheadrightarrow b$, $a \twoheadrightarrow c$ has convergent reductions for b and c at the bottom and right.[1]

Figure 5 shows that adjunction of e.d.'s at different open corners of a reduction diagram commutes. It is not hard to see that, as a consequence, the final outcome of the construction process is independent of the order of picking corners to be filled with e.d.'s. It is not independent, though, of the choice of e.d.'s to fit in, if there is a choice.

Note that we can distinghuish again, as we did with e.d.'s, the notions of *abstract* and of *geometric* reduction diagram. When drawing (geometric) reduction diagrams, we will again mostly omit the direction of the reduction arrows, which always is down or to the right—see Figure 6. Nodes always represent objects, and edges, representing reduction steps, bear an index. This information may be supplied in labels. It will become relevant to do so in Section 8.

One may think of such a geometric reduction diagram as the point set in the real plane, obtained by the union of the point sets of the geometric e.d.'s involved. (As a matter of fact two point sets: that of the edges and that of the nodes.)

4.2 Infinite reduction diagrams

Infinite reduction diagrams arise if the process of tiling with elementary reduction diagrams does not terminate, i.e., when at each finite stage open corners remain. Now we can take an infinite reduction diagram to be the union of the reduction diagrams at the stages of an infinite tiling process. This makes sense in both the abstract and the geometric sense, where our notion of limit is just the union of point sets in the plane. This way the result (or limit) of a tiling process always exists. The limit is either finite and completed, or infinite.

A familiar example of an infinite reduction diagram is the Escher diagram in Figure 1. It arises from the ARS in Figure 7, the figure also illustrating the cyclic process that leads to the infinite diagram.

Just as in the finite case, a *completed* infinite reduction diagram is defined as one having no open corners. However, observe that in contrast to the finite case, a completed infinite reduction diagram with initial reductions $a \twoheadrightarrow b$ and $a \twoheadrightarrow c$ does *not* yield converging reductions for b and c.

It is important to note that in the infinite case a limit diagram may still not be completed; namely, when a certain open corner, that has to be filled in, is forever neglected in the diagram construction.

[1] The stepwise construction of reduction diagrams is somehow reminiscent of the process of completion via proof reduction, cf. Bachmair et al. [1].

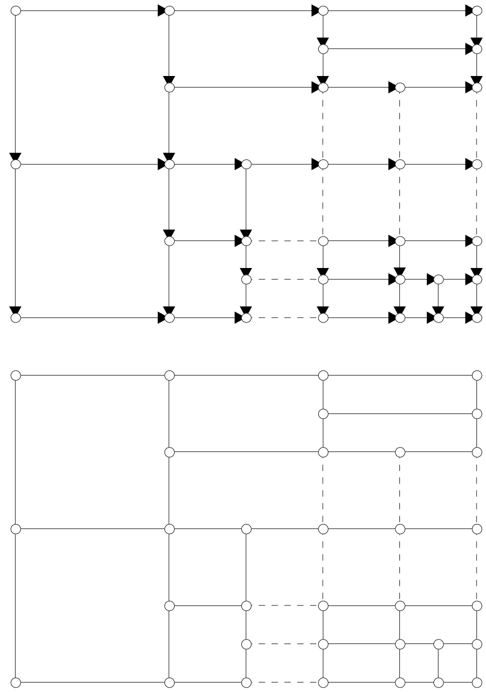

Figure 6: Completed reduction diagrams

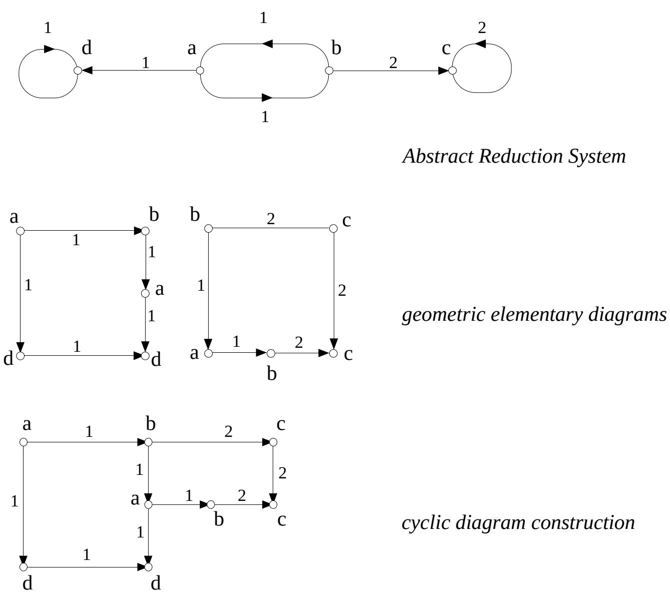

Figure 7: Cyclic construction of Escher diagram

In two ways completed infinite diagrams can always be obtained, however. The first is to allow transfinite tilings. By elementary set-theoretic considerations it follows that regardless of a strategy a completed diagram can always be obtained by transfinitely prolonging the tiling procedure. This way a completed diagram would even be obtained by a transfinite random tiling. The second is by following a fair tiling strategy, not persistently forgetting any open corners. It is not difficult to design such a fair strategy. By Corollary 4 each transfinite construction can be 'compressed' to one that has length ω.

Definition 2

1. Two tiling processes, starting with the same initial reductions, will be called *compatible* if at the same open corner always the same e.d. is adjoined. Here open corners correspond in two different tiling constructions if they have the same geometric position, to be measured, e.g., relative to the initial node of the whole diagram.

2. A, possibly transfinite, tiling strategy is called *complete* if in the limit it leads to a completed reduction diagram.

The following proposition says that a completed diagram for two given initial reductions is unique, that is, independent of the order in which open corners are filled, as long as one takes care that at the same open corner always the same e.d. is adjoined.

Proposition 3 Given two initial reductions, and two complete tiling strategies that are compatible, the resulting completed reduction diagram is unique.

Proof Let a diagram be ordered below another, if the latter arises from the former by adjoining e.d.'s. By uniqueness of filling open corners, this is easily seen to yield a lattice. It can be *normally* completed into $\mathcal{D}$ in a way preserving least upper bounds (see e.g. Davey & Priestley [4]). By the Knaster-Tarski fixed point theorem any monotone operation on $\mathcal{D}$ has a fixed point. Since adjunction of an e.d. at a given open corner is easily defined on $\mathcal{D}$ and seen to be a monotone operation, it follows that any tiling strategy has a fixed point. We conclude by remarking that a diagram is a fixed point of a complete tiling strategy iff it has no open corners. This is only the case for the top of the lattice. Hence all fixed points and all completed reduction diagrams reached by any tiling strategy are identical. □

An alternative way of formalizing all this would be to use the theory of ω-algebraic cpo's, as developed in Plotkin [13]. Since there are only countably many finite diagrams, their completion will be an ω-algebraic cpo. As a consequence all infinite diagrams are the limit of an ω-chain of finite diagrams. This could make for an interesting case study in infinite diagram construction. Anyhow, for the moment we need only the conception of an infinite diagram as a plane figure, being the limit of a tiling process. This should be sufficiently clear by now.

Corollary 4 (compression) The result of any transfinite tiling process leading to a completed diagram could also be obtained in ω many steps.

Proof Since the adoption of a fair strategy leads to a completed diagram in at most ω many steps, this follows from the last proposition. □

Remark 5 This compression result leads to a simple but fundamental observation concerning infinite reduction diagrams. Each node (or edge or e.d., for that matter), since it occurs at a finite stage of a diagram construction of length ω, has a finite 'history'. A consequence is that, as illustrated in Figure 8, above an elementary diagram (the shaded rectangular zone in Figure 8) there can not occur a condensation point of the diagram. The elementary diagram together with the zone above it must be part of some finite stage of the infinite reduction diagram.

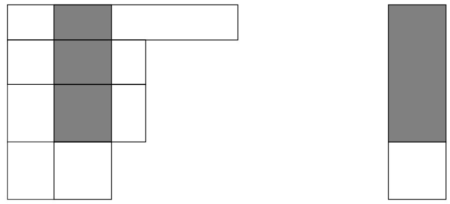

Figure 8: Part of the reduction diagram above an e.d.

The simplest infinite reduction diagram is the Escher diagram in Figure 1. Some more examples are given in Figures 9, 10 and 11. Note the fractal-like boundaries that arise in Figure 10. In each example, the diagram construction involves a certain recursion that is not hard to read off from the drawings.

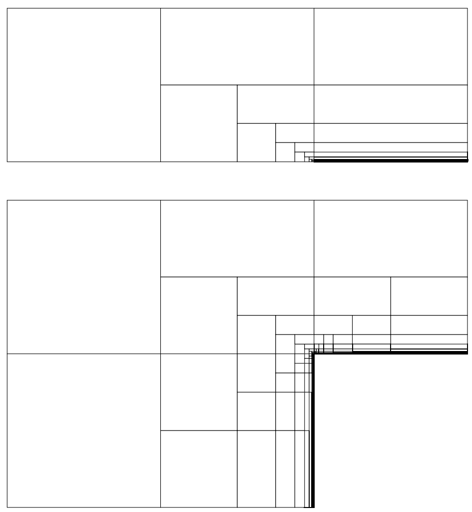

Figure 9: Infinite reduction diagrams

Remark 6 Since we admit also empty steps, it is not immediately clear that an infinite diagram contains infinitely many non-empty edges. However, this is indeed the case; Bezem et al. [2] proves the stronger fact that an infinite diagram possesses an infinite reduction containing infinitely many splitting steps. (An elementary diagram is *splitting* if one of the converging sides contains two or more steps which then are called splitting steps. Recall that by point 7 of the characterization of geometric e.d.'s in Section 3 splitting steps are always non-empty.)

5. Towers in infinite reduction diagrams

A notion that will be needed in our analysis of infinite reduction diagrams is that of a *tower*. Roughly, a tower in an infinite reduction diagram is the result of adjoining elementary diagrams in a linear way, as suggested in Figure 12. Towers are either horizontal or vertical. These notions are dual, so we need only to define horizontal towers. We will only be interested in infinite towers.

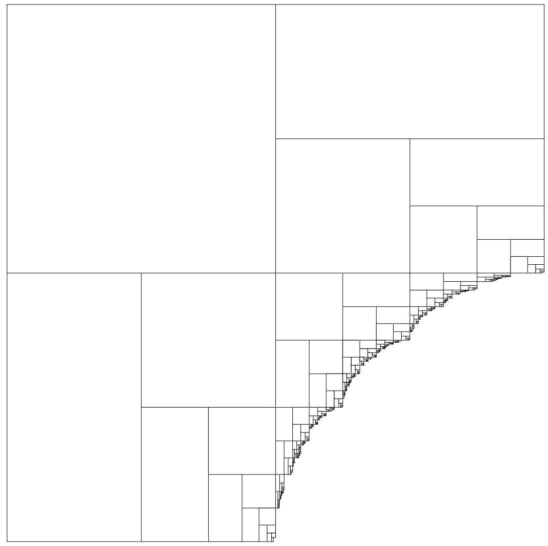

Figure 10: Infinite reduction diagram with fractal-like boundary

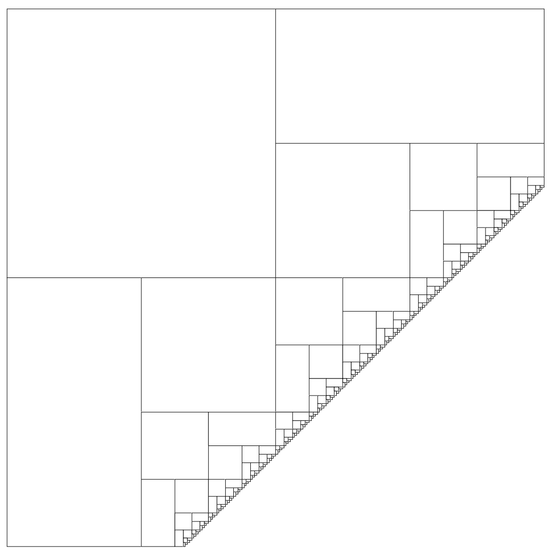

Figure 11: Infinite diagram construction with diagonal border

Definition 7 Consider an infinite reduction diagram $\mathcal{D}$, completed or not. A *horizontal tower* T is a conglomerate (or: the union) of a countably infinite set $E = \{\mathcal{E}_1, \mathcal{E}_2, \dots\}$ of e.d.'s that are already present in $\mathcal{D}$, satisfying the following two conditions:

1. The left side of $\mathcal{E}_1$ is one of the initial reduction steps of $\mathcal{D}$.
2. For each $n \geq 1$ the left initial side of E_{n+1} coincides with one of the edges on the right (converging) side of E_n.

A *vertical tower* is defined dually.

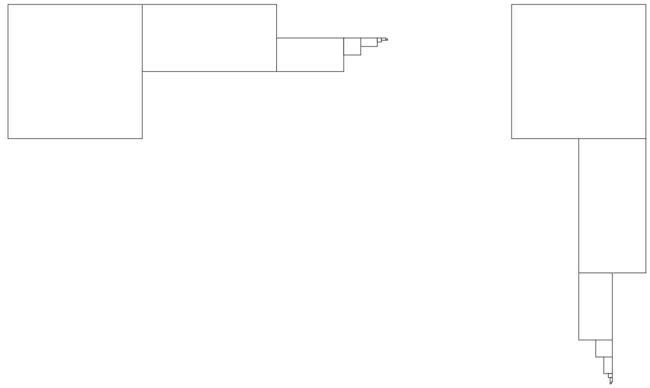

Figure 12: Horizontal and vertical tower

Figure 13 displays two towers in the first fractal-like diagram of Figure 10; Figure 14 displays the two towers constituting the Escher diagram of Figure 1. The horizontal tower is shaded, the vertical tower is blank.

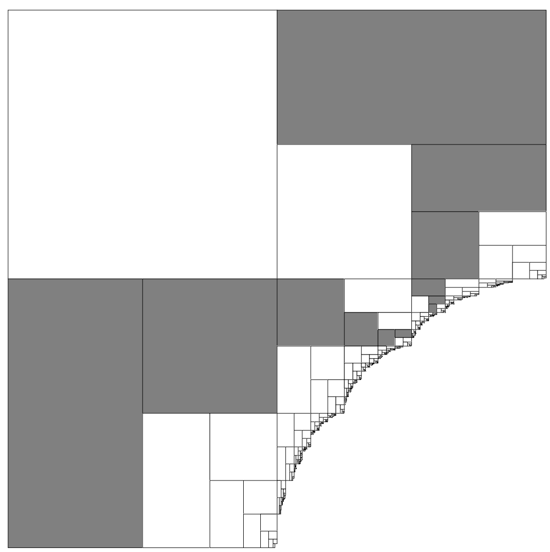

Figure 13: Towers in infinite reduction diagram

Proposition 8 Every infinite diagram contains an infinite horizontal tower and an infinite vertical tower.

Proof Consider the infinite diagram, and draw in each tile arrows from the left side to the steps in the right side (see Figure 15). In this way finitely many trees arise. By the pigeon-hole principle and König's Lemma, one of these trees must have an infinite branch. This branch determines an infinite horizontal tower. Dually we find an infinite vertical tower. □

Consider again the left-to-right trees in the preceding proof. Their branches are linearly ordered according to whether the one is 'above' the other. A branch s is *above* branch t, when after running together for some (possibly 0) steps, s branches off to above compared to t.

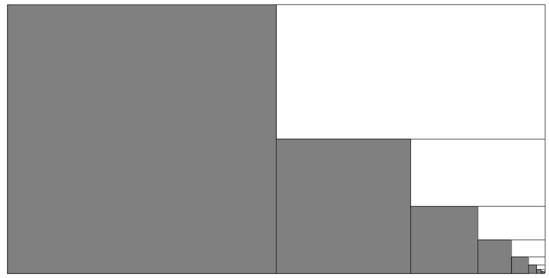

Figure 14: Towers in Escher diagram

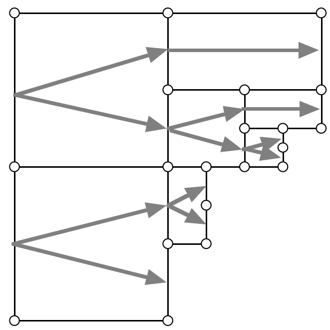

Figure 15: Finding an infinite horizontal tower

Furthermore it is clear that there is a highest infinite branch in the left-to-right trees of an infinite diagram. It is constructed in the obvious way: to start, choose the highest root of the left-right trees that has an infinite branch, then choose the highest successor with the same property, and so on.

Since branches in the left-right trees correspond with horizontal towers, there also exists a highest horizontal infinite tower. (And a leftmost vertical tower, for that matter.) This will play an important rôle later on.

Remark 9 In fact, the horizontal towers of a reduction diagram are linearly ordered by the relation 'above'. There may be continuum many towers. For example in Figure 16 there are continuum many vertical towers.

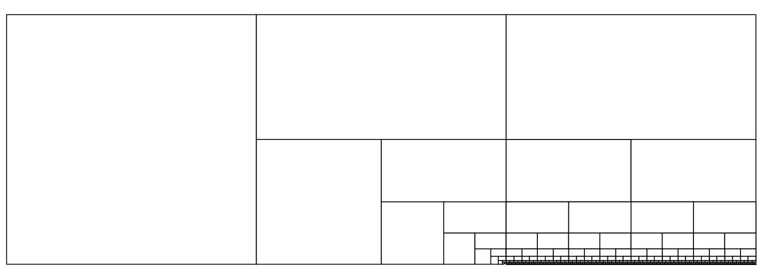

Figure 16: Continuum many towers

6. Tree coverings of reduction diagrams

In this section we define the concept of a *tree covering* of a reduction diagram. Tree coverings are the result of composing *tracing patterns* in the elementary diagrams in the reduction diagram.

Definition 10 Given an e.d. $\mathcal{E}$, a *tracing pattern* P for $\mathcal{E}$ is a collection of arrows leading from initial edges of $\mathcal{E}$ to the converging edges. The pattern P has to be such that:

1. For each edge on a converging side there is precisely one arrow leading to it.
2. An arrow leading to an empty side originates in the opposite initial side.

We say that an edge on a converging side is *traced back*, by backwards following an arrow, to one of the two initial edges.

So each converging edge can be traced back uniquely according to a given pattern. By contrast, an initial edge can trace forward to several converging edges, to one, or even to none. Examples of tracing patterns are given in Figure 17.

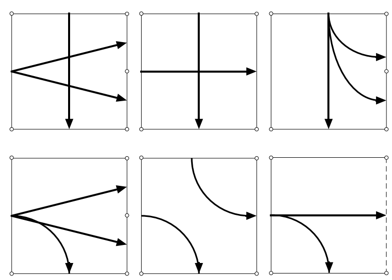

Figure 17: Elementary diagrams with tracing patterns

If all e.d.'s constituting a reduction diagram $\mathcal{D}$ are equipped with a tracing pattern, then in a natural way a pattern of *branches* emerges, by composing the ingoing and the outgoing arrows. We call such a pattern a *tree covering*. It is important to note that by following the branches backwards, each edge, anywhere in $\mathcal{D}$, can be traced back uniquely to one of the initial edges of $\mathcal{D}$.

Figure 18 shows an example of a finite, completed reduction diagram with a tree covering. Observe that the branches of the trees may intersect. Figure 19 contains a number of 'periodic' tree coverings

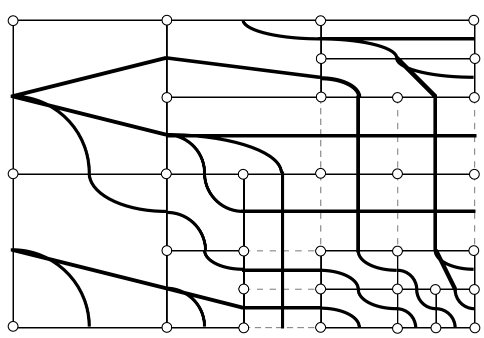

Figure 18: A tree covering

of the Escher diagram. The upper part of Figure 19 gives some of the possible tracing patterns (not exhaustive) of the elementary diagram of which the Escher diagram is built. (Note that the Escher diagram is indeed built from e.d.'s of a single shape.) These e.d.'s with trace patterns are then used to build the Escher diagram in various combinations 11, 12, etc. For example 23 means that the e.d. with trace pattern 2 is used, next the e.d. with trace pattern 3 (after mirroring); then the 23 configuration is recursively repeated.

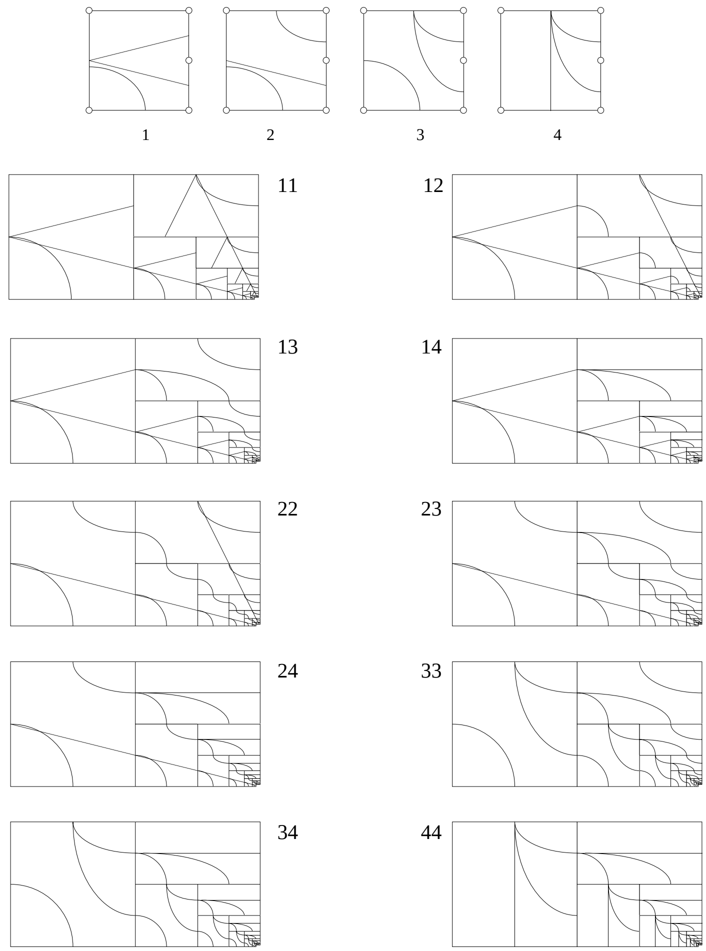

Figure 19: Periodic tree coverings

Definition 11

1. An arrow in a branch is *straight* if it leads from an initial edge to an opposing edge.
2. A branch B *changes orientation* in an e.d. $\mathcal{E}$, if it enters $\mathcal{E}$ on a vertical edge, and exits at a horizontal edge, or vice versa (equivalently, if B's arrow in $\mathcal{E}$ is not straight).
3. An infinite branch is *meandering* if it changes orientation infinitely often.
4. An infinite branch is *eventually straight* if it is not meandering. That is, if—possibly after some initial meandering— all its constituting arrows are straight.

5. Let t, s be branches that are concurrent for some steps but separate in the e.d. $\mathcal{E}$, where t is straight and horizontal. Then we say that t *branches off downward* at $\mathcal{E}$ to branch s, if s leaves $\mathcal{E}$ at a lower opposing edge than t does, or if s changes orientation. Dually we define t *branches off to the right* at $\mathcal{E}$ to s.

Observe that in each reduction diagram there is exactly one tree covering all of whose steps are straight. We call it the *canonical tree covering*. An example is given by Figure 20.

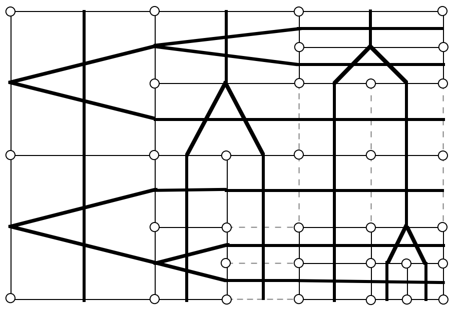

Figure 20: A reduction diagram with its canonical tree covering

Remark 12 Obviously, for an eventually straight branch there is always a tower that eventually contains it. Note that conversely an infinite horizontal tower does not always eventually contain an eventually infinite straight branch; see e.g. in Figure 19 the tree covering 34.

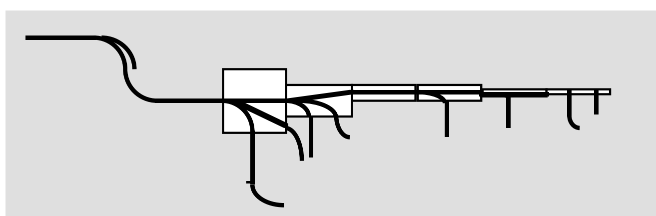

Figure 21: Infinite branch, eventually in a horizontal tower, branching only downward

Now consider an infinite reduction diagram, with a tree covering, and an infinite horizontal tower in it. Consider of each elementary diagram in the tower, its upper edge (see the heavy edges in Figure 22). Trace back each of these upper edges all the way to the initial diverging reductions of the diagram. Then, by a simple argument using the fact that the covering trees in the diagram are finitely branching, at least one infinite branch arises that we will call an *upper boundary branch* of the tower under consideration. It has the property that from any point on it infinitely many upper edges of the tower are reachable (by some branch of the tree covering).

The following definition and propositions formalize this account of the construction of an upper boundary branch.

Definition 13 Let $\mathcal{D}$ be an arbitrary infinite reduction diagram, with a tree covering, and let T be a horizontal tower in $\mathcal{D}$.

1. By an *upper edge* of T we mean an upper edge of one of the e.d.'s in T.

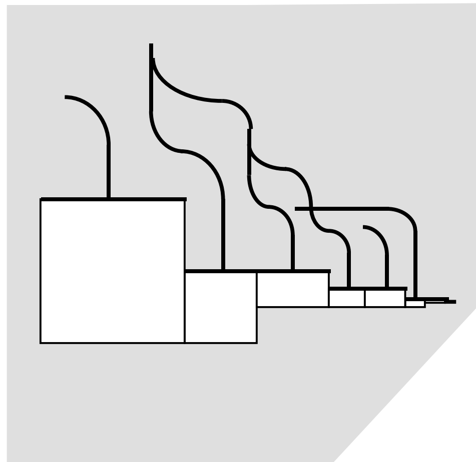

Figure 22: Upper edges and upper edge branches

2. Each upper edge of T can be traced, along a branch of the given tree covering, all the way back to the initial diverging reductions of the diagram $\mathcal{D}$. Such a path from an upper edge back to one of the initial edges of $\mathcal{D}$ is called an *upper edge branch* of T.

3. An *upper boundary branch* of T is an infinite branch s of the tree covering, such that each initial segment of s coincides with an initial segment of an upper edge branch of T.

Note that an upper boundary branch is itself not an upper edge branch, since the latter are all finite. Figure 22 shows upper edge branches. Figure 23 gives an example of an infinite horizontal tower with upper boundary branch, unique in this case; note that it is not eventually straight. This configuration can actually be found in the periodic tree coverings 22, 23, 24, 33, 34 and 44 of Figure 19. Figure 24

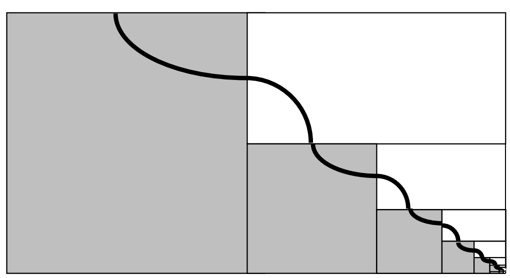

Figure 23: An upper boundary branch

gives an example of an infinite horizontal tower with exactly two upper boundary branches, one of them straight, the other one not.

Proposition 14 Every horizontal tower T has an upper boundary branch.

Proof Consider the upper edge branches of T. Since there are infinitely many of these, and since there are only finitely many initial edges of $\mathcal{D}$, by pigeon holing at least one initial edge will be the

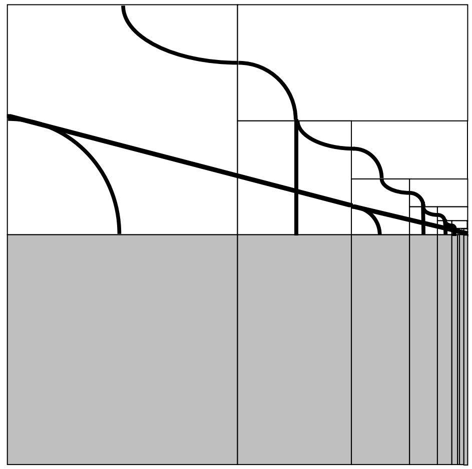

Figure 24: Horizontal tower with two upper boundary branches

origin of infinitely many upper edge branches. Choose such an initial edge and consider the infinite tree formed by all the upper edge branches originating from that edge. Since this is a finitely branching infinite tree, by König's Lemma it must have an infinite branch, say s. We claim that s will be an upper boundary branch of T.

To prove the claim we must show that each initial segment of s is also the initial segment of infinitely many upper edge branches. So consider an initial segment s_1 of s. By the construction of s as a union of upper edge branches, s_1 can be extended to an upper edge branch e_1. Since any upper edge branch is finite, there must be a (first) further point on s that is not on e_1. Consider the initial segment s_2 corresponding to that further point, and repeat the construction, resulting in a second upper edge branch e_2 and a still further point on the upper boundary branch. Continuing this process indefinitely yields infinitely many upper edge branches $e_1, e_2, \ldots$ that all extend the original initial segment s_1. □

Corollary 15 From any point on an upper boundary branch of T infinitely many upper edges of the tower T are reachable.

Proof Consider the initial segment corresponding to a point P on the upper boundary branch. By Definition 3 there are infinitely many upper edge branches that extend it. All end points of these upper edge branches are reachable from P and are on an upper edge of T. □

7. Impossible tree coverings

Now, getting to the heart of the argument of this paper, we demonstrate the impossibility of certain tree coverings for infinite reduction diagrams. In Theorem 16 three properties of tree coverings of infinite diagrams are listed, and it is proved that no tree covering can have all these three properties together. It is instructive to consider the ten cases of Figure 19. For each of these cases Table 1 sums up which of the properties 1-3 of theorem 16 are satisfied. Indeed, no case has all three properties.

Theorem 16 An infinite reduction diagram does not possess a tree covering such that:

1. All infinite branches are eventually contained in towers.

2. Infinite branches eventually contained in horizontal towers split, eventually, only downwards.

	(i)	(ii)	(iii)
11	+	-	-
12	+	-	+
13	+	-	+
14	+	-	+
22	-	+	+
23	-	+	+
24	-	+	+
33	-	+	+
34	-	+	+
44	-	+	+

Table 1: Properties of tree coverings

3. Infinite branches eventually contained in vertical towers split, eventually, only to the right.

Proof For a proof by contradiction, assume that $\mathcal{D}$ is an infinite reduction diagram with a tree covering satisfying clauses 1-3. Consider in $\mathcal{D}$ the highest infinite horizontal tower T. Let s be an upper boundary branch of T; by proposition 14 we know that there must exist one. By clause 1 the branch s must be eventually contained in a tower T', which may be horizontal or vertical.

Case 1. T' is horizontal. Since T is the highest horizontal tower, T' must be T or be lower than T. Both cases are contradictory, since by the second clause s can branch off (after some steps) only in downward direction, hence can never be a boundary branch of T.

Case 2. T' is vertical. This requires more argument to show its impossibility. Consider the relative position of the towers T, T'; there are three possibilities, captured in Figures 25, 26, 27.

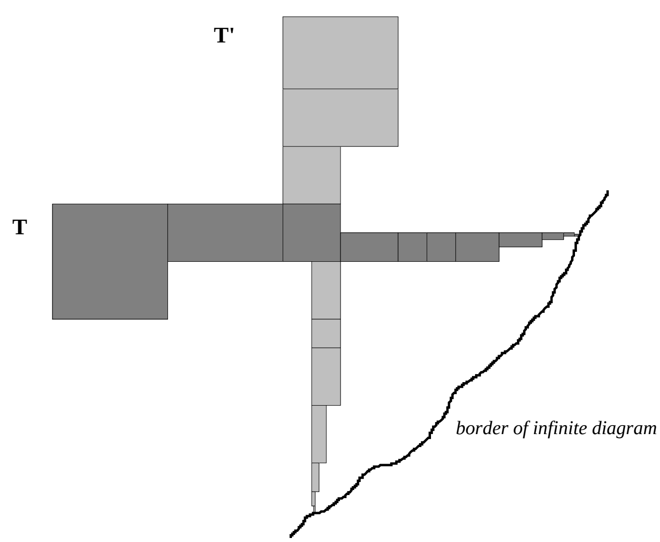

Figure 25: Relative positions: intersecting towers

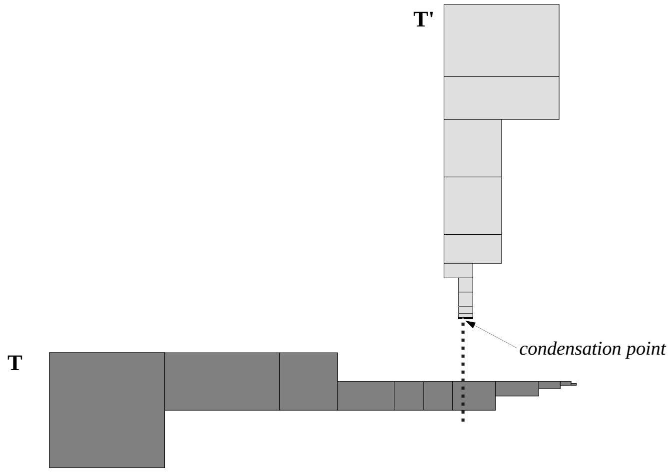

Figure 26: Relative positions: passing without intersection

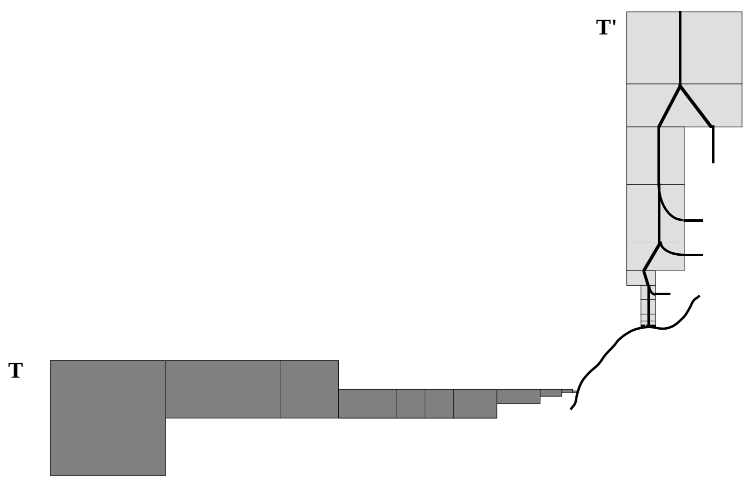

Figure 27: Relative positions: no passing, no intersection

- The first case, Figure 25, where the vertical tower T' intersects the horizontal tower T, is impossible: the part of the upper boundary branch s contained in T' below the intersection can never reach the upper edges of T.
- Figure 26, where the horizontal tower T proceeds beyond the vertical line starting at a condensation point of T', is equally impossible. This situation would contradict the observation in Remark 5.

- So only the case of Figure 27 remains as possibility. But in this case the branch s contained in tower T', can not reach more than finitely many upper edges of T, since eventually s branches off only to the right. This contradicts Corollary 15.

□

8. Confluence by decreasing diagrams

De Bruijn [5] gave a very strong confluence criterion for abstract reduction systems with indexed reduction relations. It consists of a combinatorial property of the distribution of indexes in the elementary diagrams. The original formulation in de Bruijn [5] was asymmetrical; van Oostrom [12, 11] gave a symmetrical version, as follows.

Definition 17 Define an elementary diagram to be *decreasing*, if it has the form as shown in Figure 28. This means that given two diverging steps $a \to_n b$ and $a \to_m c$ with indices n, m there is a common

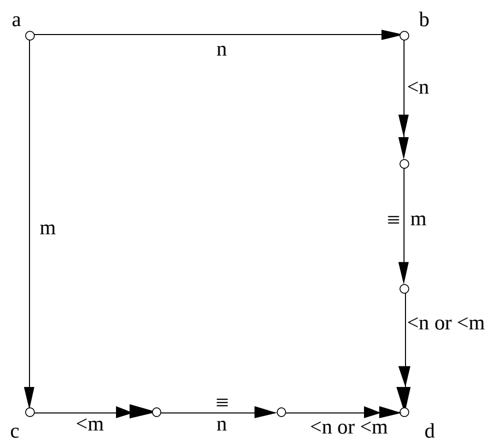

Figure 28: Elementary decreasing diagram

reduct d such that

$$b \twoheadrightarrow_{<n} \cdot \to_m^{\equiv} \cdot \twoheadrightarrow_{<n \text{ or } <m} d, \text{ and}$$

$$c \twoheadrightarrow_{<n} \cdot \to_n^{\equiv} \cdot \twoheadrightarrow_{<n \text{ or } <m} d.$$

So from b we take finitely many (possibly zero) steps with indices $< n$, followed by zero or one steps with index m, followed by some steps with index $< n$ or $< m$, with result d. Dually, from c we have a reduction to d as indicated.

In Figure 29(a) some non-decreasing elementary diagrams are given; in (b) some decreasing elementary diagrams. (The labels are subject to the usual ordering $<$ on natural numbers.)

We will now connect the present definition with the tree coverings of above, by supplying decreasing diagrams with tracing patterns. Recall that according to Definition 10, such a pattern traces the converging steps back to the two initial, diverging steps. In doing so, it will be helpful to use a heavy arrow in case the index remains the same, and a light arrow in case the index decreases.

The heavy and light arrows are determined as follows. Consider the vertical reduction $b \twoheadrightarrow_{<n} \cdot \to_m^{\equiv} \cdot \twoheadrightarrow_{<n \text{ or } <m} d$. If this reduction is empty, then the right side of the ensuing e.d. is an empty step, which we trace back with a heavy arrow to the vertical initial step $a \to c$. If it is not empty we do the following:

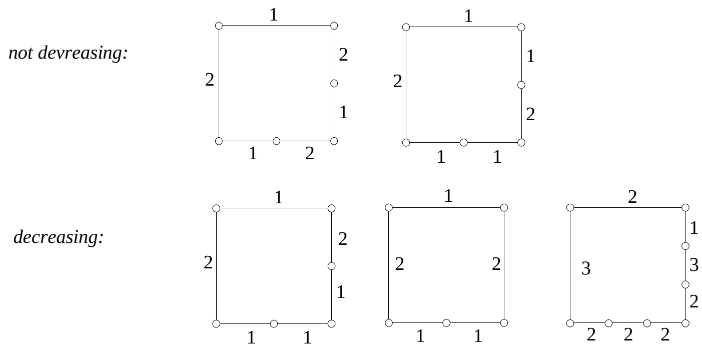

Figure 29: Decreasing and non-decreasing elementary diagrams

- We let the first part of this reduction, consisting of steps with index less than the index n of the horizontal step $a \to_n b$, trace back lightly to that step.
- If the second part consists of one step with label m, it is traced back heavily to the vertical step
- The part consisting of steps with label less than n or m is treated as follows. If the step label is less than n we trace back lightly to $a \to b$, if less than m then lightly to $a \to c$, if both then we choose one.

Likewise dually.

So a decreasing elementary diagram with the tracing arrows has one of the shapes of Figure 30: containing two heavy arrows, or one, or none. It is important that heavy arrows (along which the indices remain the same) are straight, while the light arrows (along which the indices decrease) may involve a change of orientation. See Figure 31, consisting of the decreasing elementary diagrams of

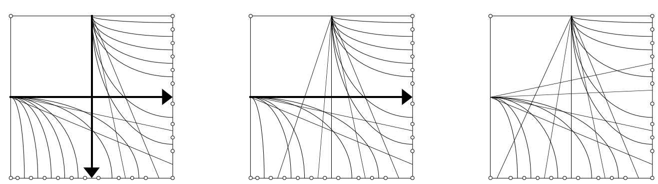

Figure 30: Elementary diagrams with tree covering

Figure 29 but now enriched with the tracing arrows (with the convention for heavy and light just mentioned). Note that the tracing pattern (the tree covering) is not uniquely determined by the decreasing elementary diagram; e.g. Figure 32 contains two tracings for the same elementary diagram.

We now have the following proposition.

Proposition 18 Every diagram construction using decreasing elementary diagrams will terminate eventually in a finite confluent diagram.

Proof Equip the decreasing elementary diagrams with heavy and light arrows as explained above. Note that heavy arrows preserve indices and are straight, while light ones decrease indices and may change orientation. Note furthermore that a horizontal heavy arrow cannot split off in upward direction (see Figure 30) and likewise dually.

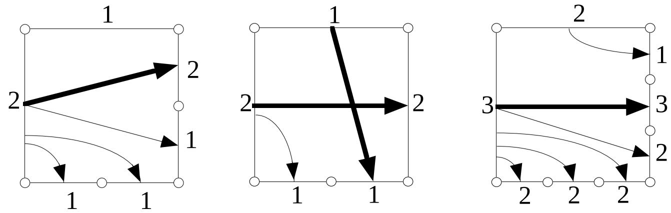

Figure 31: Decreasing diagrams of Figure 29 with tracing arrows

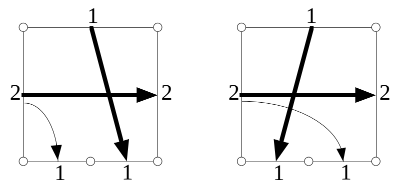

Figure 32: One elementary decreasing diagram with alternative tracings

Now consider an infinite branch in the diagram enriched with heavy and light arrows. Because the partial order I is well-founded, eventually only heavy (index-preserving) arrows can occur in this branch. But these are straight. So, every infinite branch must be eventually straight (and thus contained in a tower).

Furthermore, from infinite horizontal branches we can eventually only have split offs in downward direction (either by straight light arrows, or by a change in orientation, see Figure 30). Likewise dually. That is, the three hypotheses of Theorem 16 are fulfilled. According to this theorem the diagram cannot be infinite. □

Corollary 19 (Confluence by decreasing diagrams) Every ARS with reduction relations indexed by a well-founded partial order I, and satisfying the decreasing criterion for its elementary diagrams, is confluent.

9. Eliminating empty steps

Tiling with elementary diagrams, the way we did it in this paper, and as it was introduced in Klop [9], involves the use of improper e.d.'s in order to cope with empty steps. In van Oostrom [12, p. 30] it was noted that empty steps can be avoided, by passing from a rewrite relation $\rightarrow$ to its reflexive closure $\rightarrow^{\equiv}$. Indeed, to prove confluence of an ARS is equivalent to proving confluence of its reflexive closure. Even stronger, it is easy to see that an ARS has decreasing diagrams if and only if its reflexive closure has decreasing diagrams with only non-empty converging sides (if a side would be empty, then a 'reflexive step' can be inserted). It is not hard to see that working with the reflexive closure would yield *exactly the same* constructions as now in the paper; technically there would be no difference.

One could say allegorically that we have introduced empty steps 'at run time' (making them more tangible and understandable) while they could have been introduced 'at compile time' (making for more efficient and compact code).

Remark 20 The construction above implies that removing the reflexive closures in the converging sides in the elementary decreasing diagrams as defined by Figure 28, would not decrease the power of the decreasing diagrams theorem, Corollary 19.

Acknowledgements
We would like to thank Stefan Blom and Jan Rutten for useful discussions.

References

1. L. Bachmair, N. Dershowitz, and J. Hsiang. Orderings for equational proofs. In *Proceedings of the 1st IEEE Symposium on Logic in Computer Science*, pages 346–357, Cambridge MA, 1986.
2. Marc Bezem, Jan Willem Klop, and Vincent van Oostrom. Diagram techniques for confluence. *Inform. and Comput.*, 141(2):172–204, 1998.
3. Mirna Bognar. A point version of decreasing diagrams. In Joeri Engelfriet and Tigran Spaan, editors, *Proceedings Accolade'96*, pages 1–14, Amsterdam, 1997. Dutch Graduate School in Logic.
4. B.A. Davey and H.A. Priestley. *Introduction to lattices and order.* Cambridge University Press, 1990.
5. N.G. de Bruijn. A note on weak diamond properties. Memorandum 78-08, Eindhoven University of Technology, August 1978.
6. J.R. Hindley. *The Church-Rosser property and a result in combinatory logic.* PhD thesis, University of Newcastle-upon-Tyne, 1964.
7. G. Huet. Confluent reductions: Abstract properties and applications to term rewriting systems. *Journal of the ACM*, 27(4):797–821, 1980.
8. J. W. Klop. Term rewriting systems. In D. Gabbay S. Abramski and T. Maibaum, editors, *Handbook of Logic in Computer Science*, volume 1. Oxford University Press, 1992.
9. J.W. Klop. *Combinatory Reduction Systems.* Mathematical Centre Tracts Nr. 127. CWI, Amsterdam, 1980. PhD Thesis.
10. M.H.A. Newman. On theories with a combinatorial definition of 'equivalence'. *Annals of Mathematics*, 43(2):223–243, 1942.
11. V. van Oostrom. Confluence by decreasing diagrams. *Theoretical Computer Science*, 126:259–280, 1994.
12. V. van Oostrom. *Confluence for Abstract and Higher-Order Rewriting.* PhD thesis, Vrije Universiteit, Amsterdam, March 1994.
13. G.D. Plotkin. Post-graduate lecture notes in advanced domain theory (incorporating the "Pisa notes"). Department of Computer Science, University of Edinburgh, 1981.
14. J. Staples. Church-Rosser theorems for replacement systems. In J. Crossley, editor, *Algebra and Logic*, number 450 in Lecture Notes in Mathematics, pages 291–307. Springer-Verlag, 1975.