Acta Informatica (2013) 50:123–156
DOI 10.1007/s00236-012-0174-y

ORIGINAL ARTICLE

Infinitary rewriting: closure operators, equivalences and models

Stefan Kahrs

Received: 15 June 2011 / Accepted: 19 November 2012 / Published online: 4 December 2012

Abstract Infinitary Term Rewriting allows to express infinite terms and transfinite reductions that converge to those terms. Underpinning the machinery of infinitary rewriting are closure operators on relations that facilitate the formation of transfinite reductions and transfinite equivalence proofs. The literature on infinitary rewriting has largely neglected to single out such closure operators, leaving them implicit in definitions of (transfinite) rewrite reductions, or equivalence relations. This paper unpicks some of those definitions, extracting the underlying closure principles used, as well as constructing alternative operators that lead to alternative notions of reduction and equivalence. A consequence of this unpicking is an insight into the abstraction level at which these operators can be defined. Some of the material in this paper already appeared in Kahrs (2010). The paper also generalises the notion of equational model for infinitary rewriting. This leads to semantics-based notions of equivalence that tie in with the equivalences constructed from the closure operators.

1 Introduction

Infinitary rewriting deals with infinite terms, which are typically defined through the metric completion of finite terms through some metric. In the simplest case (metric d_∞) this is equivalent [5,23] to a co-inductive definition of terms, i.e. the set of *infinitary terms* $Ter^\infty(\Sigma)$ is the *largest set* such that every t in this set has some root symbol F taken from the signature Σ and n direct subterms t_i $(1 \leq i \leq n)$ that are all in $Ter^\infty(\Sigma)$, where n is the arity of F as defined by the signature. In other words, infinitary terms are defined co-inductively through the way they unfold, without a guarantee that this unfolding ever comes to an end. *Infinite* terms are indeed those where it does not.

Metric completion is a general-purpose semantic construction on metric spaces which "adds" to a metric space limits to all its Cauchy-sequences. More accurately, there is always a dense isometric embedding of the original space into a complete metric space.

S. Kahrs (✉)
Department of Computer Science, University of Kent at Canterbury, Canterbury CT2 7NF, UK
e-mail: S.M.Kahrs@kent.ac.uk

Using metric completion with other metrics on terms than d_∞ can restrict the infinite terms under consideration [16], as $Ter^\infty(\Sigma)$ can also be seen as a final co-algebra [19]. As we shall see later, equational reasoning is affected by the presence or absence of infinite terms in the term universe.

The complexities of infinitary rewriting arise in (and as a result of) the constructions of relations between infinite terms. Infinity arises here in two ways: the objects of concern can be infinite, but so are the reductions and proofs that link them: for instance, building an infinite term step-by-step through rewriting should take infinitely many reduction steps, as does modifying an infinite term in infinitely many different places. This means that relations on infinitary terms need to come with a machinery that supports such reasoning in the infinite.

Much of this machinery can be explained through (monotonic) maps on relations (on terms): if R is some relation on terms, then $X(R)$ could be another relation on terms that can apply R repeatedly in some manner. In finite abstract rewriting, the most commonly used maps X of this kind are the transitive closure and the equivalence closure, and in term rewriting the rewrite closure (closure under contexts and substitutions). Because they are closure operators there is also no point repeating the construction: $X(R) = X(X(R))$. In this paper, we adopt this point of view for infinitary rewriting: we extract the relational maps implicit in some known constructions, and investigate them: are they closure operators? If not, what would they be? Can we lift the maps to a higher level of abstraction? Does the semantics suggest alternatives?

Regarding transfinite rewriting itself, the traditional [8] construction is to first generalise reduction sequences to transfinite ones, and then to add limits to those sequences. The result of this construction is the notion of (weakly) convergent reduction $\twoheadrightarrow_w$: we have $t \twoheadrightarrow_w u$ iff there is a (possibly infinite) continuous reduction sequence that starts at t and ends in u. This construction can be seen as $W(\rightarrow_R) = \twoheadrightarrow_w$, i.e. there is a map W on relations which produces $\twoheadrightarrow_w$ when given the single-step rewrite relation $\rightarrow_R$ as input. In practice, $\twoheadrightarrow_w$ has turned out to be difficult to work with. One of the problems with $\twoheadrightarrow_w$ is that it is in general not closed under its own construction principle: $W(W(\rightarrow_R)) \neq W(\rightarrow_R)$. A traditional alternative [20] to this is the notion of strongly continuous reduction $\twoheadrightarrow_s$ which constrains the reduction sequences to which limits should be added. Its standard definition involves the concept of rewrite rule and therefore is based on the notion of a rewrite system, not just a relation; as we shall see this definition can be tweaked: there is a monotonic map S on relations such that $S(\rightarrow_R) = \twoheadrightarrow_s$; moreover, S is a closure operator on rewrite relations.

Beyond that, we are looking at non-traditional constructions that add limits more directly to relations, using the topological closure of a set as its primary tool. Of those, the pointwise closure operator $P^\star$ shows particular promise with various similarities to S.

Regarding equivalences, they have so far largely been ignored in infinitary rewriting. Perhaps the main reason is that the standard way of forming them is as an equivalence closure of some other relation. However, an equivalence closure does not incorporate any form of limit-taking and is therefore a very weak form of closure when we need to reason with infinite terms.

One equivalence that has been looked at in the literature is $\sim_{hc}$, the "equivalence modulo hypercollapsing terms" [20]. This relation is meant to identify all meaningless terms, similar to a sensible λ-theory in the λ-calculus [4]. Hyper-collapsing terms are in orthogonal iTRSs similar to the unsolvable terms of λ-calculus—we cannot extract information out of them. The original reason why this notion was invented was to patch up a "misbehaviour" of orthogonal iTRS in the presence of collapsing rules—rules whose right-hand sides are variables.

Technically, $\sim_{hc}$ is applying an equivalence relation $\sim$ across a potentially infinite term, where $\sim$ relates all hyper-collapsing terms to one another. One could describe this closure principle as a generalisation of substitutivity to potentially infinite terms: $(\forall x \in Var.\theta(x) \sim \sigma(x)) \Rightarrow \theta(t) \sim \sigma(u)$. This principle can be derived from a semantic construction.

The concept of equational model [7–9,18,36] is tightly connected to the notion of equivalence, since each class of models entails an equivalence of its own: $t =_R u$ iff the equation $t = u$ holds in *all equational models* in the class R. For an equation $t = u$ to hold in a model A means that $[\![t]\!]_A^\rho = [\![u]\!]_A^\rho$ is true for all valuations $\rho : Var \to A$, where $[\![\_]\!]_A$ is the interpretation map for A, providing for the interpretation of (possibly) non-ground terms in A. Thus the notion of model is tied to the notion of an interpretation map, and we gain extra generality by treating the interpretation map as a primitive notion, rather than a derived one.

Mainly for this purpose we generalise (and to some extent: correct) the notion of equational model, basing it on the concept of Eilenberg-Moore algebra from Category Theory [27]—which also uses interpretation maps as a primitive notion.

Some of the material presented here already appeared in [18]. This applies to Sects. 3.1, 3.3, 4.1 (although proofs have been changed and results generalised), and Sect. 4.2. The treatment of models and equivalence is here substantially different, and the concept of a semi-deciding rewrite relation is not discussed here at all.

2 Preliminaries

We need several notions from Topology [11,15,25] and Infinitary Rewriting. This section contains those definitions, to make the paper self-contained.

We call a property P on subsets of a set A *closable* if the intersection $\bigcap_{i\in I} A_i$ of any family of subsets $A_i \subseteq A, i \in I$ that each satisfy P itself satisfies the property. As I may be empty, A has to satisfy P too. If P is closable we can form the P-closure of any subset $K \subseteq A$, the smallest subset of A that contains K and satisfies P.

One particular way of creating a closable property is through a monotonic function $F : \wp(A) \to \wp(A)$, as being a post-fixpoint of F is a closable property: if $X_i, i \in I$ is a family of subsets of A then $F(\bigcap_{i\in I} X_i) \subseteq \bigcap_{i\in I} F(X_i)$ by monotonicity of F; if in addition $\forall i \in I.\ F(X_i) \subseteq X_i$ (each X_i is a post-fixpoint of F) then $\bigcap_{i\in I} F(X_i) \subseteq \bigcap_{i\in I} X_i$ by monotonicity of intersection, and therefore $\bigcap_{i\in I} X_i$ is a post-fixpoint of F too. We write $F^\star(X)$ for the smallest post-fixpoint of F containing X which by an instance of the Knaster-Tarski fixpoint theorem [6, Theorem 4.11] is also a fixpoint of the function $G(Y) = F(Y) \cup X$.

In particular, this concept will be used on relations, viewing them as sets of pairs. Clearly, the conjunction of closable properties is closable.

2.1 Topology

A *topological space* is a pair $(S, \mathcal{O})$ where S is a set and $\mathcal{O}$ a subset of $\wp(S)$ such that it is closed under finite intersections and arbitrary unions, and $\emptyset, S \in \mathcal{O}$. $\mathcal{O}$ is also called a topology over S. The elements of $\mathcal{O}$ are called *open* sets, their complements w.r.t. S are called *closed* sets. The *closure* of a subset $A \subset S$ is the intersection of all closed sets that contain A, and is written $\mathrm{Cl}(A)$. A *neighbourhood* of $x \in S$ is a set $N \subseteq S$ for which there is an $A \in \mathcal{O}$ such that $x \in A \subseteq N$. A point $z \in S$ is called *discrete* iff $\{z\}$ is open.

Given a set S a *subbase* is a subset of $\wp(S)$. The topology generated from $A \subset \wp(S)$ is the smallest topology over S (ordered by set inclusion) that contains A. A subbase is called a *base* iff its elements are closed under binary intersection.

A topological space S is called T_0 iff for any two different points $x, y \in S$ there is an open set $A \subset S$ that contains the one point but not the other. S is called T_1 iff all its singleton sets are closed. It is T_2 (or *Hausdorff*) iff any two distinct points in S have disjoint neighbourhoods. Further commonly used separation properties are T_4 (and T_3) where closed sets can be separated from disjoint closed sets (for T_3: singletons) by open neighbourhoods.

For a binary relation $R \subset A \times B$, the function $R^{-1} : \wp(B) \to \wp(A)$ is defined as usual as $R^{-1}(X) = \{a \in A \mid \exists x \in X.\, a\ R\ x\}$, and similarly for functions. A function $f : A \to B$ between topological spaces is called *continuous* iff $f^{-1}(X)$ is open whenever X is open. A function $f : A \to B$ between topological spaces is an *open map* iff it preserves open sets; it is a *closed map* iff it preserves closed sets. A relation R between topological spaces A and B is called *lower semi-continuous* (or lsc) if R^{-1} preserves open sets, and *upper semi-continuous* (or usc) if R^{-1} preserves closed sets.

Two topological spaces are called *homeomorphic* iff they are connected by an *isomorphism*, i.e. a bijection that is continuous in both directions.

A subset F of a topological space is called *compact* iff whenever $F \subseteq \bigcup_{i \in I} A_i$, where each A_i is open, then there is a finite subset $J \subseteq I$ such that $F \subseteq \bigcup_{i \in J} A_i$.

Given a set A and a topological space B, the set of set-theoretic functions $A \to B$ can be equipped with the product topology [15]—generated from the subbase that each function set $\{f \mid f(a) \in O\}$ is open for any $a \in A$ and any open subset O of B. In this topology, sequences of functions converge to another function iff they do so pointwise. However, if A is also a topological space and a sequence of continuous functions converges in $A \to B$ its limit is not necessarily continuous [12, 28.1.1]. We use this topology to equip valuations or substitutions with a topology.

A *metric space* is a set (M, d) where M is a set and $d : M \times M \to \mathcal{R}$ a distance function such that for all $x, y, z \in M$: (i) $d(x, y) = 0 \iff x = y$ and (ii) $d(x, z) \leq d(x, y) + d(z, y)$. This axiomatisation implies $d(x, y) \geq 0$ and $d(x, y) = d(y, x)$. For an *ultra-metric space* (ii) is replaced by the stronger condition $d(x, z) \leq \max(d(x, y), d(z, y))$. The topology of a metric space is defined as follows: $A \subseteq M$ is open iff $\forall x \in A.\exists \epsilon > 0.\forall y \in M.\, d(x, y) < \epsilon \Rightarrow y \in A$. A topological space is called *metrizable* iff it is homeomorphic to a metric space.

The *metric completion* of (M, d) is the unique (up to isomorphism) metric space $(M^\bullet, d^\bullet)$, with a function $e : M \to M^\bullet$, such that e preserves distances and $e(M)$ is *dense* in $M^\bullet$, i.e. $\mathrm{Cl}(e(M)) = M^\bullet$.

A function between metric spaces A and B is *uniformly continuous* iff $\forall \epsilon > 0.\ \exists \delta > 0.\ \forall x, x' \in A.\, d_A(x, x') < \delta \Rightarrow d_B(f(x), f(x')) < \epsilon$. Uniformly continuous functions have unique continuous extensions to the respective metric completions. A special case are *non-expansive* functions where $\delta = \epsilon$.

A function f on the interval $[0, 1]$ is called *amenable* if $f(x) = 0 \iff x = 0$. A (unary) *ultra-metric map* (short: umm) is an amenable monotonic function. An n-ary function $g : [0, 1]^n \to [0, 1]$ is called a umm iff it can be expressed as $g(x_1, \ldots, x_n) = \max_i g_i(x_i)$ where each g_i is a umm — the g_i are called the *components* of g. In this paper we only consider continuous ultra-metric maps; for more details on umms consult [16].

2.2 Sequences

A *sequence* over a set T is a function $f : \alpha \to T$ where α is an ordinal—viewed as a von Neumann ordinal, i.e. the set of all strictly smaller ordinals. f is called *open* if α is a limit ordinal, otherwise it is *closed*.

If T is a topological space, then the sequence is called *continuous* if f is continuous as a function—using the standard order topology on ordinals. The standard order topology for a partially ordered set is generated from the *base* of open intervals $(a, b) = \{x \mid a < x \wedge x < b\}$

and open half-intervals $(a, \infty) = \{x \mid a < x\}$, $(-\infty, a) = \{x \mid x < a\}$, i.e. all open sets in the order topology arise as (possibly infinitary) unions of intervals and half-intervals. Consequently, any successor ordinal $\gamma + 1$ is a discrete point because $\{\gamma + 1\} = (\gamma, \gamma + 2)$.

We say that a continuous open sequence (with domain α) converges iff it can be extended to a continuous closed sequence (with domain $\alpha + 1$).

A *subsequence* of f is a strictly monotonic function $g : \beta \to \alpha$. g is called *cofinal* [13] iff $\forall \gamma < \alpha. \exists \zeta < \beta. g(\zeta) \geq \gamma$. We say that a subsequence g converges iff the sequence $f \circ g$ does. Using cofinal subsequences is a common construction in infinitary rewriting [19,31].

Notice that a subsequence of a closed sequence is cofinal iff it is itself closed and picks the final element, i.e. if $\alpha = \alpha' + 1$ then $\beta = \beta' + 1$ and $g(\beta') = \alpha'$. In particular, the shortest cofinal subsequence of a closed sequence has length 1 whilst cofinal subsequences of open sequences have at least length ω. Generally, if a sequence converges to c then so do all its cofinal subsequences.

Example 1 Consider the sequence $f : \omega + \omega \to \mathcal{R}$ defined as follows:

$$\begin{aligned} f(0) &= 1 \\ f(n+1) &= (f(n) + 2/f(n))/2, \text{ if } n < \omega \\ f(\omega + k) &= \sqrt{2} + (k \bmod 3) \end{aligned}$$

This sequence is continuous, because it is continuous at ω, the only limit ordinal in its domain. It does not converge though. The subsequence $g : \omega \to \omega + \omega$ of f defined as $g(\alpha) = 2 * \alpha$ does converge (to $\sqrt{2}$) but it is not cofinal. The subsequence $h : \omega \to \omega + \omega$ of f defined as $h(\alpha) = \omega + 1 + 3 \cdot \alpha$ also converges (to $1 + \sqrt{2}$) and it is cofinal too.

2.3 Infinitary term rewriting

The notions of signature and algebra used here are unsorted versions.

A *signature* is a pair $\Sigma = (\mathcal{F}, \#)$ where $\mathcal{F}$ is a set (of function symbols) and $\# : \mathcal{F} \to \mathcal{N}$ is the function assigning each symbol its arity. We assume a countably infinite set *Var* of *variables*, disjoint from $\mathcal{F}$. The set of *finite terms* over Σ is called $Ter(\Sigma)$ and it is defined to be the smallest set such that (i) $Var \subseteq Ter(\Sigma)$ and (ii) $F(t_1, \ldots, t_n) \in Ter(\Sigma)$ whenever $F \in \mathcal{F} \wedge \#(F) = n \wedge \{t_1, \ldots, t_n\} \subset Ter(\Sigma)$. The *root symbol* of a term $F(t_1, \ldots, t_n)$ is F, the root symbol of a variable x is x.

We write $\mathcal{N}^*$ for the free monoid over the natural numbers (i.e. finite words), with neutral element $\langle\rangle$ and infix $\cdot$ as monoid multiplication. The set of positions $Pos(t)$ of a finite term t is the smallest subset of $\mathcal{N}^*$ such that: (i) $\langle\rangle \in Pos(t)$, (ii) $\{i \cdot q \mid 1 \leq i \leq n, q \in Pos(t_i)\} \subset Pos(F(t_1, \ldots, t_n))$. Positions are used for selecting or replacing subterms. For selecting the subterm at position p of a term t we use the notation $t|_p$. For replacing the subterm at position p of a term t by u we use the notation $t[u]_p$.

A Σ-algebra is a set A together with functions $F_A : A^n \to A$ for every $F \in \mathcal{F}$ with $\#(F) = n$. A *valuation* into A is a function $\rho : Var \to A$. Any Σ-algebra A determines an interpretation function $[\![\_]\!]_A : Ter(\Sigma) \to (Var \to A) \to A$ as follows:

$$\begin{aligned} [\![x]\!]_A^\rho &= \rho(x), \qquad \text{if } x \in Var \\ [\![F(t_1, \ldots, t_n)]\!]_A^\rho &= F_A\left([\![t_1]\!]_A^\rho, \ldots, [\![t_n]\!]_A^\rho\right) \end{aligned}$$

A homomorphism between Σ-algebras A and B is a function $h : A \to B$ that commutes with interpretation: $h([\![t]\!]_A^\rho) = [\![t]\!]_B^{h \circ \rho}$. An algebra A in a class of algebras $\mathcal{K}$ is initial if, for any algebra $B \in \mathcal{K}$ there exists a unique homomorphism $\phi : A \to B$.

A substitution is a homomorphism on $Ter(\Sigma)$, viewed as a Σ-algebra, where $F_{Ter(\Sigma)}(t_1, \ldots, t_n) = F(t_1, \ldots, t_n)$. A unary context is a term $C[\ ]$ with a unique occurrence of a special variable $\Box$ which we call the hole. We write $C[t]$ for the term obtained from replacing $\Box$ with t.

Infinitary terms are defined through a metric completion process. A commonly used metric is d_∞ [1]. It can be defined inductively on finite terms: $d_\infty(t, u) = 1$ iff t and u have different root symbols; $d_\infty(t, t) = 0$; otherwise the root symbols are the same: $d_\infty(F(t_1, \ldots, t_n), F(u_1, \ldots, u_n)) = 1/2 * \max_{1 \le i \le n} d_\infty(t_i, u_i)$. $(Ter(\Sigma), d_\infty)$ is an ultra-metric space and we write $(Ter^\infty(\Sigma), d_\infty)$ for its metric completion, which is also a Σ-algebra [16, Corollary 1]. The notions of positions, subterm selection and subterm replacement carry over from finite terms [16].

More commonly, the distance $d_\infty(t, u)$ is described as 2^{-k} where k is the length of the shortest path that shows a differences between t and u. However, the equivalent inductive characterisation is an instance of a more general pattern. First defined in [16], there is a more general notion of *term metric* m from which distance functions d_m and their corresponding metric completions $Ter^m(\Sigma)$ can be derived: a term metric m is a Σ-algebra with carrier set $[0, 1]$ such that for each $F \in \mathcal{F}$ the function F_m is an ultra-metric map. Then $d_m(t, u) = 1$ if t and u have different root symbols, and otherwise

$$d_m(F(t_1, \ldots, t_n), F(u_1, \ldots, u_n)) = F_m(d_m(t_1, u_1), \ldots, d_m(t_n, u_n)).$$

Here it becomes clear that such d_m functions are homomorphisms of Σ-algebras: the product $A \times B$ of any two Σ-algebras A and B is itself a Σ-algebra with $F_{A\times B}((a_1, b_1), \ldots, (a_n, b_n)) = (F_A(a_1, \ldots, a_n), F_B(b_1, \ldots, b_n))$, so in particular this applies to $A = B = Ter^m(\Sigma)$. A *metric signature* Σ_m is a pair (Σ, m) where Σ is a signature and m a term metric for Σ.

For each unary context C and term metric m there is a unary ultra-metric map C_m defined inductively on the hole position: $\Box_m = id$; otherwise, C has a root symbol F, and for some $i \in \omega$, $C|_i = D$ where D is a unary context. Then $C_m = F_{m,i} \circ D_m$ where $F_{m,i}$ is the i-component of F_m. The function C_m computes distances: if $d_m(t, u) = \epsilon$ then $d_m(C[t], C[u]) = C_m(\epsilon)$ [16]. In the special case of term metric ∞ we get $C_\infty(\epsilon) = 2^{-|p|} \cdot \epsilon$, where $|p|$ is the length of the position p of the hole in C.

Each function symbol F is uniformly continuous on $Ter^m(\Sigma)$ [16, Proposition 5]; moreover:

Proposition 1 *Each function symbol is an open map on* $Ter^m(\Sigma)$.

Proof Consider any open sets of terms $T_1, \ldots, T_n$ and the set $T = \{F(p) \mid p \in T_1 \times \cdots \times T_n\}$. We need to show that T is open, i.e. that for every $t \in T$ there is an $\epsilon > 0$ such that all terms within ϵ-distance of t are also in T. Consider the term $F(t_1, \ldots, t_n) \in T$. Because all T_i are open there are distances $\epsilon_i > 0$ such that everything within ϵ_i-distance of t_i is also in T_i. Let $F_{m,i}$ be the components of the ultra-metric map F_m. We can set $\epsilon = \min_{1 \le i \le n} F_{m,i}(\epsilon_i)$. If $d_m(t, u) < \epsilon \le 1$ then u must be of the form $F(u_1, \ldots, u_n)$. Thus $d_m(t, u) = F_m(d_m(t_1, u_1), \ldots, d_m(t_n, u_n))$. By [16, Lemma 1] we generally have $F_m(a_1, \ldots, a_n) < \min_i F_{m,i}(b_i) \Rightarrow \forall j.\, a_j < b_j$, so this implies $d_m(t_i, u_i) < \epsilon_i$ for all i, hence $u_i \in T_i$ for all i and $u \in T$. □

The concepts of substitutions and contexts extend uniquely from the finite into the infinitary setting [16]. A relation R on terms is called *substitutive* iff $t\ R\ u \Rightarrow \sigma(t)\ R\ \sigma(u)$ for any substitution σ; it is called *compatible* if $t\ R\ u \Rightarrow C[t]\ R\ C[u]$ for any context $C[\ ]$. A relation is called a *rewrite relation* iff it is both compatible and substitutive.

A rewrite rule is pair (l, r), usually written $l \to r$, such that $l, r \in \mathit{Ter}^m(\Sigma)$ (for some term metric m), $l \notin \mathit{Var}$, and all variables occurring in r also occur in l. A rewrite rule $l \to r$ is called *collapsing* if $r \in \mathit{Var}$, and *left-linear* if l is a *linear* term, i.e. if variables occur in it at most once. A term l is called a *pattern* iff it is finite and linear.

Note: the standard definition of rewrite rule for infinitary TRS restricts left-hand sides to be finite [22]. However, the main motivation behind this restriction is to make matching decidable. To truly achieve this we need in addition that l is linear, because matching against non-linear terms requires equality checks, and the equality of infinite terms is undecidable. Thus if we are concerned about decidability, l should be finite and linear, otherwise neither restriction is necessary.

Also both restrictions are necessary for the compression of strong continuous reduction sequences [20], and for the following property:

Lemma 1 *Consider substitution application, viewed as a function* $\mathit{subst} : \mathit{Ter}^m(\Sigma) \to (\mathit{Var} \to \mathit{Ter}^m(\Sigma)) \to \mathit{Ter}^m(\Sigma)$*, where* $\mathit{Var} \to \mathit{Ter}^m(\Sigma)$ *takes the product topology, Then each* $\mathit{subst}(t)$ *is continuous and it is also open if and only if* t *is a pattern.*

Proof Continuity was shown in [16, Proposition 2]. That $\mathit{subst}(t)$ is an open map for patterns can be shown by induction on the structure of t. If $t \in \mathit{Var}$ then $\mathit{subst}(t)$ is a projection from the product and thus open. Otherwise $t = F(a_1, \ldots, a_n)$ where each a_i is a pattern. Applying an open set of substitutions to each a_i gives by induction hypothesis open sets of terms A_i. Thus the (finite) product of these A_i is open, and as t is linear it is equal to the substitution instances of the tuple $(a_1, \ldots, a_n)$. Finally, recall that F is an open map by Proposition 1.

That $\mathit{subst}(t)$ is not an open map if t is not a pattern is easy to see: if t is infinite then so are all its substitution instances, and non-empty sets containing only infinite terms are never open, because metric completion is a dense embedding. If t is non-linear, e.g. $t = C[x, x]$ then $\theta(t) = \theta(C)[\theta(x), \theta(x)]$. If $\theta(x)$ is an infinite term then for any sequence of finite terms u_n converging to $\theta(x)$ the sequence $a_n = \theta(C)[\theta(x), u_n]$ is converging to $\theta(t)$, but none of the a_i is a substitution instance of t. □

Another open map for term manipulation is context application. We write $\mathit{Cont}^m(\Sigma)$ for the set of unary contexts, and use the metric d_m on it to equip it with a topology—treating the hole □ as a distinct variable. Context application can then be expressed as a function

$$\mathit{cont} : \big(\mathit{Cont}^m(\Sigma) \times \mathit{Ter}^m(\Sigma)\big) \to \mathit{Ter}^m(\Sigma),$$

with $\mathit{cont}(C, t) = C[t]$.

Note: to establish that context application is an open map the proof uses a further notion from [16]: the equivalence relation $\sim$ on unary contexts; $C \sim D$ means that C and D have the hole in the same position p and the same function symbols at prefixes of p.

Lemma 2 *Context application is an open and continuous map.*

Proof Continuity of context application is another consequence of [16, Proposition 2].

Let A be an open set of contexts, T be an open set of terms, and $C \in A,\ t \in T$. We need to show that $\mathit{cont}(A \times T)$ is a neighbourhood for $C[t]$. As A and T are open, for some $\epsilon, \delta > 0$ any context within ϵ-distance of C is in A and any term within δ-distance of t is in T; we need to find a $\gamma > 0$ such that any term v within γ-distance of $C[t]$ takes the form $v = D[u]$ where $D \in A$ and $u \in T$. We can set $\gamma = \min(\epsilon, C_m(\delta))$ where C_m is the umm associated to the context C. As umms are amenable we have $C_m(\delta) > 0$ and therefore $\gamma > 0$.

If $d_m(v, C[t]) < C_m(\delta)$ then by [16, Lemma 2] there is a context D and term u such that $v = D[u]$ and $d_m(u, t) < \delta$ and $D \sim C$. Therefore $u \in T$. Contexts in the relationship $\sim$

have the hole in the same position and therefore $d_m(C, D) = d_m(C[x], D[x])$. Using [16, Proposition 6 (iii)] we get overall:

$$d_m(C, D) = d_m(C[x], D[x]) \le d_m(C[t], D[u]) = d_m(C[t], v) < \gamma \le \epsilon$$

Therefore $D \in A$ and $v \in \mathit{cont}(A \times T)$. □

An iTRS is a pair (Σ_m, R) where Σ_m is a metric signature and R a set of rewrite rules for that signature. The relation $\to_R$ is the smallest rewrite relation containing R as a subrelation.

This means: if $t \to_R u$ then there is a context $C[\,]$, a substitution σ and a rule $l \to r$ such that $t = C[\sigma(l)]$ and $u = C[\sigma(r)]$.

Lemma 3 *Let $l, r \in \mathit{Ter}^m(\Sigma)$. If l is a pattern then the substitutive closure of the relation $\{(l, r)\}$ is lsc.*

Proof Let U be any open set of terms. Consider the set of substitutions $\Theta = \{\theta \mid \theta(r) \in U\}$ which is open (in the product space $\mathit{Var} \to \mathit{Ter}^m(\Sigma)$), because $\Theta = \mathit{subst}(r)^{-1}(U)$ and $\mathit{subst}(r)$ is continuous (Lemma 1). The set $L = \{\theta(l) \mid \theta \in \Theta\}$ is the inverse image of U w.r.t. the rule, and it must be open by Lemma 1. □

In particular, the lemma applies to rewrite rules of that shape, but it is slightly more general, allowing l to be a variable or r to contain arbitrary variables.

Proposition 2 *The compatible closure of any lsc relation (on $\mathit{Ter}^m(\Sigma)$) is lsc.*

Proof Let R be an lsc relation on $\mathit{Ter}^m(\Sigma)$, and let S be the compatible closure of R. Let $U \subseteq \mathit{Ter}^m(\Sigma)$ be open. We need to show that $S^{-1}(U)$ is open. Because context application is continuous (Lemma 2), $\mathit{cont}^{-1}(U)$ is an open set of pairs (C, t) where C is a context and t a term. Projecting these from $\mathit{cont}^{-1}(U)$ gives an open set of contexts A and an open set of terms T. Because R is lsc $R^{-1}(T)$ is open and as context application is an open map (Lemma 2 again), we have that $\mathit{cont}(A \times R^{-1}(T))$ is open. Expressed pointwise, that set is $\{D[u] \mid D[t] \in U \wedge u\ R\ t\}$, which is the same as $S^{-1}(U)$. □

Proposition 3 *If all rules have patterns as left-hand sides then $\to_R$ is lsc.*

Proof $\to_R$ can be constructed as the union of the compatible closures of the substitutive closures of all rules.

The substitutive closure of a rule is lsc by Lemma 3. This is preserved by the compatible closure by Proposition 2. And lsc relations are preserved under arbitrary union [25]. □

Notice that the proof of Proposition 3 ultimately only hinges on that (i) substitution application and context application are both continuous, (ii) that substitution application to patterns is an open map, and (iii) that context application is an open map. Thus, notions of infinitary rewriting that use other topologies share this result as long as they have these (topological) properties—and along the way the notions of patterns and evaluation contexts could even be modified.

3 Transfinite sequences

In the following we are looking at several notions of defining transfinite relations over some topological space T. Some principles can be defined at this abstract level alone, because convergence is a topological concept, but others are less abstract and even specific to iTRSs.

If we consider a term universe that is based on partial orders such as in [2,23] we can derive for that a topology by first turning the order into a pre-CPO using ideal completion [6, Lemma 3.20] and then equip this structure with the Scott-topology [32]—the Alexandrov-topology would be "too discrete" to be of interest for infinitary rewriting as it does not permit sequences of finite terms to converge to infinite ones. Since $Ter^{\infty}(\Sigma)$ can itself be viewed as the result of an ideal completion [23] the first step is typically redundant. Thus these constructions provide terms with a notion of convergence, even though these topologies are not metrizable as they are not T_1. As we shall see later they even provide a general notion of *strong convergence*, for any topology on terms.

Term universes that are merely restrictions on infinite terms, such as Isihara's proper terms [14], are always metrizable, though not necessarily metrizable as a complete metric space.

Generally we are forming transfinite reduction relations $X(R)$, where R is some (otherwise given) binary relation on our topological space T and X a function on relations over T which in some way accounts for repeatedly applying R. For iTRSs, R is typically the single-step rewrite relation and X some closure principle.

All the functions X we consider are monotonic and thus give rise to closure operators $X^{\star}$, which are the least fixpoints of the function $G(R) = X(R) \cup R$. This fixpoint also arises as the limit of the sequence $Y_0 = \emptyset$, $Y_{n+1} = G(Y_n)$, $Y_{\lambda} = \bigcup_{\beta<\lambda} Y_{\beta}$, where the indexing set for the sequence Y could be any ordinal that well-orders the domain of the relations. This is so because relations on a set form a CPO, so [6, Theorem 4.14] applies.

Generally, for a monotonic function X on relations to be a closure operator we need [6, Definition 2.20] that it is idempotent ($X(X(R)) = X(R)$) and that it is increasing, $R \subseteq X(R)$.

When there is no ambiguity, we take $\rightarrow_R$ to be clear from the context and write $\twoheadrightarrow_x$ for $X(\rightarrow_R)$ and $\twoheadrightarrow_{xx}$ for $X(X(\rightarrow_R))$, especially when X may fail to be a closure operator itself. This notation does not imply that $\twoheadrightarrow_x$ is automatically a pre-order, although it often will be.

We want functions that add limits to be closure operators, as this is an indication of the stability of the concept; it also has technical applications, e.g. the property $X(\rightarrow_R) = X(X(\rightarrow_R))$ features in the proofs of [17]. Even more importantly is that we can define equivalence relations that are closed under X, and these are by construction closed under $X^{\star}$ too.

Reduction sequences for a relation $R \subseteq T \times T$ are those where neighbouring elements are within the single-step reduction relation R, i.e. if $f : \alpha \rightarrow T$ is our reduction sequence then $f(n)\ R\ f(n+1)$, provided $n + 1 < \alpha$.

This construction works fine for *finite* sequences. For infinite sequences this definition fails to put any constraints whatsoever what happens at $f(\lambda)$, for limit ordinals λ.

3.1 Standard solution: weak convergence

The traditional choice to fix this problem is to demand that the sequence be continuous, which means that the indexing function f is continuous. In the standard order topology, for a sequence f to be continuous at λ, $f(\lambda)$ must be the (unique) limit of $f(\gamma)$, as γ approaches λ from below.

Closed sequences give us a way of defining the relation $\twoheadrightarrow_w$: $t \twoheadrightarrow_w u$ iff there is some ordinal α and some closed continuous reduction sequence $f : \alpha+1 \rightarrow T$ such that $f(0) = t$ and $f(\alpha) = u$. The reason closed sequences are used for this is two-fold: (i) it applies to both finite and transfinite sequences, and (ii) they have a designated end point. Open sequences may not converge, but even if they do they may converge to a choice of limits (in non-Hausdorff spaces).

In the infinitary rewriting literature the relation $\twoheadrightarrow_w$ is often called "weak convergence" [20]. This concept immediately generalises to sequences over an arbitrary topological space T.

Remark 1 The definition asks for "some ordinal" without giving an upper bound which could be regarded as problematic. However, we can give an upper bound for the length of continuous sequences on a metric space M: ω_1, the first uncountable ordinal. The main reason is that any continuous function $f : \omega_1 \to M$ is eventually constant. In [33, p. 70] this is shown for real-valued functions, but their proof immediately generalises to arbitrary metric spaces as it is only based on the distance function of real numbers.

On T_1 spaces we can also limit the length of continuous sequences. This follows from the following property: in every continuous closed sequence on a T_1 space every point has a last occurrence.

Lemma 4 *Let T be a T_1 space and $f : \alpha + 1 \to T$ be a continuous closed sequence on T. Then $\forall \gamma \leq \alpha.\, \exists \zeta \leq \alpha.\ f(\gamma) = f(\zeta) \wedge \forall \kappa \leq \alpha. f(\kappa) = f(\gamma) \Rightarrow \kappa \leq \zeta$.*

Proof Pick an arbitrary $\gamma \leq \alpha$ and consider the sequence of ordinals $\gamma_i \leq \alpha$ such that $f(\gamma_i) = f(\gamma)$. This sequence has a limit ζ (which may be α itself) in the domain of f, because f is closed. Because f is continuous at ζ the point $f(\gamma)$ occurs in every neighbourhood of $f(\zeta)$ which implies $f(\zeta) \in \mathrm{Cl}(\{f(\gamma)\})$. Since T is T_1 the set $\mathrm{Cl}(\{f(\gamma)\})$ must be a singleton and therefore $f(\gamma) = f(\zeta)$. □

Based on this one can use the (ordinal corresponding to the) cardinality of T as a natural upper bound for the length of continuous sequences we need to consider:

Proposition 4 *For T_1 spaces: if $t \twoheadrightarrow_w u$ then this is witnessed by a reduction sequence f which is an injective function.*

Proof The proof uses the well-ordering principle, and goes by induction on a well-founded total order $>$ on T. Let $g : \alpha+1 \to T$ be *any* witness for $t \twoheadrightarrow_w u$. We show that for all $b \in T$ there is an ordinal $\beta \leq \alpha$ and a reduction sequence $f : \beta + 1 \to T$ witnessing $t \twoheadrightarrow_w u$ such that whenever $f(\gamma) = f(\kappa) \leq b$ then $\gamma = \kappa$.

Let a be an element of T. By induction hypothesis we assume the result for all $b < a$. Thus there is $f' : \beta' + 1 \to T$ witnessing $t \twoheadrightarrow_w u$ such that if $f'(\gamma) = f'(\kappa) < a$ then $\gamma = \kappa$.; If a occurs more than once in the range of f' then it has a first occurrence $f'(\gamma) = a$ and, by Lemma 4, a last occurrence $f'(\zeta) = a$. Let π be the unique ordinal such that $\beta' = \zeta + \pi$ (see [29, Theorem 6.3.5]). We can set $\beta = \gamma + \pi$ and define $f(x) = f'(x)$ for $x < \gamma$ and $f(\gamma + y) = f'(\zeta + y)$. □

For non-T_1 spaces the result can fail, because we may not have last occurrences of elements of T even in the shortest witnessing reduction sequence:

Example 2 Take the set $T = \{0, 1\}$ with the trivial topology $\mathcal{O} = \{\emptyset, \{0, 1\}\}$ and the relation $0 \to 0$. Then $0 \twoheadrightarrow_w 1$ but the shortest reduction sequence witnessing that is of length ω and therefore not injective.

Corollary 1 *For any metric space T, if $t \twoheadrightarrow_w u$ then this is witnessed by reduction sequence of length less than ω_1.*

Proof Apply Proposition 4. The length of that injective sequence has to be less than ω_1, because otherwise its restriction to domain ω_1 would remain continuous, therefore eventually constant (see Remark 1), and therefore non-injective. □

A subtle distinction between continuity and (weak) convergence (see e.g. [20]) can be made like this: a continuous sequence $f : \alpha \to T$ is called *convergent* if it is closed or if it can be extended to a continuous closed sequence $f' : \alpha + 1 \to T$.

For relations on an arbitrary topological space we can view $\twoheadrightarrow_w$ through a function W that maps a "single-step" relation to $\twoheadrightarrow_w$. Its most general type is $W : \wp(T \times T) \to \wp(T \times T)$, where T is (the underlying set of) a topological space: thus $(t, u) \in W(R)$ iff there is an ordinal α and a continuous function $f : \alpha + 1 \to T$ such that $f(0) = t$, $f(\alpha) = u$ and $\forall \gamma < \alpha.\ f(\gamma)\ R\ f(\gamma + 1)$.

The name "weak convergence" insinuates a certain unhappiness about this construction amongst researchers. This was mostly due to difficulties establishing positive properties about continuous reduction sequences in general, and the relation $\twoheadrightarrow_w$ in particular. One sign that W is difficult to work with is the following "deficiency":

Proposition 5 *For any topological space, W is monotonic, but it is in general not a closure operator.*

Proof Monotonicity of W is obvious from the construction, since if $R \subseteq R'$ then each reduction sequence over R is also a reduction sequence over R'. Here is a counterexample that shows that W is not a closure operator, because $\twoheadrightarrow_w$ and $\twoheadrightarrow_{ww}$ differ:

Example 3 Consider the iTRS with rules

$$F(A, x) \to F(B, D(x))$$
$$B \to A$$

We have $F(A, x) \to_R^2 F(A, D(x))$, but $F(A, x) \twoheadrightarrow_w F(A, D^\infty)$ does not hold, because A would change to B at every other step and thus all subsequent terms in that reduction sequences are at distance $\frac{1}{2}$, and so the sequence cannot converge. Since $F(A, x) \to_R^2 F(A, D(x))$ we also have $F(A, x) \twoheadrightarrow_w F(A, D(x))$ and consequently $F(A, x) \twoheadrightarrow_{ww} F(A, D^\infty)$. □

Because W is monotonic there is an associated closure operator $W^\star$, which we will be looking at later. In [17,19] it was shown that $W(R) = W^\star(R)$ if R is a *converging* relation on a Hausdorff space, where "converging" means that all open continuous reduction sequences over R converge. Clearly, Example 3 is not converging.

3.2 Standard modification: strong convergence

When Kennaway et al. [20] gave the relation $\twoheadrightarrow_w$ its index w and labelled convergence/continuity with the adjective "weak", they also introduced the notion of *strong convergence* and the associated relation $\twoheadrightarrow_s$. For a reduction sequence to be strongly convergent it is required that in addition to being weakly convergent, redex positions get arbitrarily deep in sufficiently tight neighbourhoods of limit ordinals.

That wording is not a suitable definition for a *function on relations*, as it involves the concept of rewrite rule, i.e. it is not just expressed in terms of the relation alone. However, this can be mended.

Definition 1 We call $C \in Cont^m(\Sigma)$ a *witnessing context* for $t\ R\ u$ if there are terms $t', u' \in Ter^m(\Sigma)$ such that $t = C[t'] \wedge u = C[u'] \wedge t'\ R\ u'$.

Witnessing contexts always exist because $\square$ is a witnessing context. The idea behind strong convergence is that the convergence is already determined by the convergence of witnessing contexts, never mind the redexes. This can be formalised as follows:

Definition 2 A reduction sequence $f : \alpha \to Ter^m(\Sigma)$ for the relation R is called strongly continuous at $\lambda \in \alpha$ iff for every neighbourhood A of $f(\lambda)$ there is a neighbourhood N of λ such that for all $\gamma, \gamma + 1 \in N$ there is a witnessing context C for $f(\gamma)\ R\ f(\gamma + 1)$ such that $C[a] \in A$ for any $a \in Ter^m(\Sigma)$. f is called a strongly continuous reduction sequence iff it is strongly continuous everywhere.

Reduction sequences are automatically strongly continuous at any non-limit ordinal γ, because it is discrete and its neighbourhood $\{\gamma\}$ does not contain any reduction steps.

The wording of the definition of strong convergence does not mention metrics. This means that we could supply $Ter^m(\Sigma)$ with a different topology than its metric space and still have a meaningful notion of strong convergence. A case in point is the so-called Böhm-extension of infinite terms with partial terms in [2]. Equipped with the Scott-topology, Bahr's concept of p-convergence becomes an instance of this general notion of strong convergence.

The reason this notion of strong convergence is compatible with the standard notion of strong convergence for d_∞ is this: for metric spaces it suffices if the condition is satisfied for neighbourhoods A that are open balls around $f(\lambda)$. This turns the condition on $C[a]$ of the definition into $d_m(C[a], f(\lambda)) < \epsilon$ for some $\epsilon > 0$—and for d_∞ that only depends on the depth of the hole position.

Proposition 6 *A reduction sequence that is strongly continuous at λ is also continuous at λ.*

Proof Continuity of the sequence f at λ means that for any neighbourhood A of $f(\lambda)$ there is a neighbourhood N of λ such that $f(N) \subseteq A$. Let M be the neighbourhood of λ witnessing strong continuity at λ for A. Then M must contain an interval N of the form $[\eta, \lambda]$ which is an open set. Let $\gamma \in N$. If $\gamma = \lambda$ then $f(\gamma) = f(\lambda) \in A$ by definition of A. If $\gamma \neq \lambda$ then $\gamma + 1 \in N$. As $N \subseteq M$ there is a witnessing context C for $f(\gamma)\ R\ f(\gamma + 1)$ such that $C[a] \in A$ for any $a \in Ter^m(\Sigma)$. As $f(\gamma) = C[t]$ for some $t \in Ter^m(\Sigma)$ we have in particular $f(\gamma) \in A$. □

The proof of Proposition 6 does not use the metric on $Ter^m(\Sigma)$, only its topology. Thus for any topology on any $Ter^m(\Sigma)$ we have this implication that strong convergence implies convergence.

Definition 3 We write $S(R)$ (or $\twoheadrightarrow_s$) for the relation for which $t \twoheadrightarrow_s u$ iff there is a strongly continuous reduction sequence $f : \alpha + 1 \to Ter^m(\Sigma)$ for R with $f(0) = t$ and $f(\alpha) = u$.

Corollary 2 *For any topology on terms, $S(R) \subseteq W(R)$.*

Proof Follows directly from Proposition 6. □

Analogous to the introduction of weak convergence, a strongly continuous reduction sequence $f : \alpha \to Ter^m(\Sigma)$ is called *strongly convergent* if it is either closed or if it can be extended to a strongly continuous closed reduction sequence $f' : \alpha + 1 \to Ter^m(\Sigma)$.

Example 4 Consider the iTRS with the rule $F(x) \to F(G(x))$. The (unique) reduction sequence $f : \omega + 1 \to Ter^\infty(\Sigma)$ with $f(0) = F(x)$ and $f(\omega) = F(G^\infty)$ is (though continuous) not strongly continuous at ω, because all reductions are at the root position, so the only witnessing context is $\Box$. We could ignore the metric d_∞ and supply the set $Ter^\infty(\Sigma)$ with a different topology $\mathcal{O}$, and check under which circumstances this reduction sequence f would be regarded as strongly continuous. Here, f is strongly continuous at ω iff the only open neighbourhood for $F(G^\infty)$ in $\mathcal{O}$ is $Ter^\infty(\Sigma)$ itself. In that case, $F(G^\infty)$ is also the $\perp$-element in the specialisation order of the topology [32].

Proposition 7 *The function $S : \wp(Ter^m(\Sigma) \times Ter^m(\Sigma)) \to \wp(Ter^m(\Sigma) \times Ter^m(\Sigma))$ is monotonic.*

Proof If $R \subseteq R'$ and $t\ R\ u$ then all contexts witnessing $t\ R\ u$ also witness $t\ R'\ u$. Hence every strongly continuous reduction sequence of R is also a strongly continuous reduction sequence for R'. □

The fact that the super-relation may have more witnessing contexts makes it possible that reduction sequences of R that are not strong may become strong for $R' \supseteq R$:

Example 5 Add to Example 4 the rule $G(x) \to G(G(x))$. The sequence f mentioned there is now strong.

Definition 4 A context C is *safe* for an open set $A \subseteq Ter^m(\Sigma)$ iff $\forall t \in Ter^m(\Sigma).\, C[t] \in A$. For a binary relation R on $Ter^m(\Sigma)$ and an open set $B \subseteq Ter^m(\Sigma)$ we say that $t\ R\ u$ is *safe for B* iff it has a witnessing context D that is safe for B.

In order to show that the function S is a closure operator we need that strongly continuous reductions that use strongly continuous reductions over R as elementary steps can be flattened. This is indeed possible, provided R is compatible:

Lemma 5 *Let R be a compatible relation. For any strong $S(R)$ reduction sequence $f : \alpha \to Ter^m(\Sigma)$ (with $\alpha \geq 1$) there is an ordinal α' and a strong R reduction sequence $g_\alpha : \alpha' \to Ter^m(\Sigma)$ with $f(0) = g_\alpha(0)$ where f can be viewed as a cofinal subsequence of g_α. Moreover, if every step in f is safe for $S(R)$ and A then every step in g_α is safe for R and A.*

Proof The proof goes by induction on α. Thus, for each $\gamma < \alpha$ with $\gamma \geq 1$ we assume an ordinal γ' by applying the lemma to the restriction of f to domain γ.

To express f as a cofinal subsequence of g_α with $f = g_\alpha \circ f'$ we define a strictly monotone $f' : \alpha \to \alpha'$ such that $f'(\gamma) = \gamma'$ for all $\gamma \in \alpha$ with $\gamma \geq 1$ and $f'(0) = 0$.

If $\alpha = 1$ then $\alpha' = 1$ and $f = g_1$.

If $\alpha = \beta + 1 \geq 2$ then f has a last element $f(\beta)$, and the cofinality condition only means that α' is of the form $\alpha' = \beta' + 1$ and $f'(\beta) = \beta'$. We make a case distinction on β:

- If $\beta = \gamma + 1$ then $f(0) \twoheadrightarrow_{ss} f(\gamma) \twoheadrightarrow_s f(\beta)$. By induction hypothesis we have $g_\beta(0) = f(0) \twoheadrightarrow_s f(\gamma) = g_\beta(\gamma')$. Suppose the step in $f(\gamma) \twoheadrightarrow_s f(\beta)$ is safe for $S(R)$ and A. Hence, there is a context C and terms a, b such that $f(\gamma) = C[a],\ f(\beta) = C[b],\ a \twoheadrightarrow_s b$ and $C[u] \in A$ for any term u. Let $h : \zeta_\gamma + 1 \to Ter^m(\Sigma)$ be a reduction sequence witnessing $a \twoheadrightarrow_s b$. From this we can construct g_α: we set its domain as $\alpha' = \gamma' + \zeta_\gamma + 1$, and define $g_\alpha(x) = g_\gamma(x)$ if $x < \gamma'$, and $g_\alpha(\gamma' + x) = C[h(x)]$ if $x < \zeta_\gamma$. The function g_α is an R reduction sequence by compatibility of R; all its steps are safe for A, in the first part by the induction hypothesis, and in the second because they have witnessing context C.
- Otherwise β is a limit ordinal. We can construct the open sequence $g_\beta : \beta' \to Ter^m(\Sigma)$ by the induction hypothesis, and extend it to g_α by setting $g_\alpha(\beta') = f(\beta)$. If all $S(R)$ steps in f are safe for A then all R steps in g_α are safe for A as they all lie in g_β. It remains to be shown that g_α is strongly (R) continuous at β'. Since f is strongly continuous at β for any open neighbourhood B of $f(\beta)$ there is a $\xi < \beta$ such that all $S(R)$ steps in f beyond ξ are safe for B. By induction hypothesis this means that all steps in g_β beyond ξ' are safe for B too.

Otherwise α is a limit ordinal. We can set $\alpha' = \lim_{\beta\to\alpha} f'(\beta)$ and define g_α as the limit of its g_β approximations. Preservation of safety follows by induction hypothesis. The subsequence f' is by construction cofinal. □

Note: the sequence g_α does not depend on which safety requirement we make, and neither does the ordinal α'. Both do depend though on the choice of the reduction sequences h that witness each $a \twoheadrightarrow_s b$. This ambiguity could be removed by appeal to Zorn's Lemma, i.e. by picking the "smallest" such reduction sequence.

Proposition 8 *The function S is a closure operator on compatible relations, i.e.* $R \subseteq S(R) = S(S(R))$ *provided R is compatible.*

Proof $R \subseteq S(R)$ holds as the (strong) continuity constraint on any finite sequence is redundant. Regarding the fixpoint property, we clearly have $\twoheadrightarrow_s \subseteq \twoheadrightarrow_{ss}$, so we only need to show the converse inclusion. However, if $t \twoheadrightarrow_{ss} u$ then we can take the witnessing strong $S(R)$ reduction sequence $f : \alpha + 1 \to Ter^m(\Sigma)$, apply Lemma 5 to obtain the strong R sequence $g : \alpha' + 1 \to Ter^m(\Sigma)$ which in itself witnesses $t \twoheadrightarrow_s u$. □

The proofs of Proposition 8 and Lemma 5 are purely topological and thus do not depend on the topology applied to $Ter^m(\Sigma)$: strongly converging reductions can always be flattened.

Since the rewrite relations of iTRSs are compatible by definition, Proposition 8 always applies to their rewrite relations $\to_R$ and hence $\twoheadrightarrow_s = \twoheadrightarrow_{ss}$ for all iTRSs.

3.3 Adherence

Proposition 5 showed that W is not a closure operator. As W is monotonic it does induce a closure operator $W^\star$. As we shall see, there is an alternative way of defining it on metric spaces.

Instead of asking for convergence we can ask for adherence: $t \twoheadrightarrow_a u$ is defined like $t \twoheadrightarrow_w u$, except for one thing: instead of being continuous the witnessing indexing function f merely needs to be "adherent": This is in a certain sense a concept dual to convergence: a converging sequence stays eventually always within a neighbourhood of a limit, an adherent sequence instead always eventually visits it. This is similar to the acceptance condition for successful runs in Büchi automata [34], where a successful run (an ω sequence) visits a final state infinitely often.

Formally, for an ordinal α and a topological space T: a function $f : \alpha \to T$ is *adherent at* $x < \alpha$ iff for all intervals $[\gamma, x)$ with $x \in \mathrm{Cl}([\gamma, x))$ we have that $f(x) \in \mathrm{Cl}(f[\gamma, x))$. The function is called *adherent* if it is adherent at every point in its domain.

In the special case $\alpha = \omega + 1$ (the simplest non-trivial scenario) adherence of $f : \omega + 1 \to T$ means that the sequence f visits every neighbourhood of $f(\omega)$ infinitely often. If T is moreover a metric space then for every $\epsilon > 0, k \in \omega$ there is an $n > k$ with $d(f(n), f(\omega)) < \epsilon$.

As for continuous sequences, we can limit α to at most ω_1 if T is a metric space: if we can extend the adherent open sequence $f : \omega_1 \to T$ with an accumulation point a then there is an initial open segment of f that also has a as an accumulation point: there must be a cofinal subsequence $g : \omega \to \omega_1$ of f such that $f \circ g$ converges to a. As g is a countable increasing sequence it itself converges to $\alpha < \omega_1$, and thus the restriction of f to domain α is also an adherent open sequence with accumulation point a.

Functions are automatically adherent at successor ordinals (and 0), and for limit ordinals $\lambda < \alpha$ the condition means that for every neighbourhood N of $f(\lambda)$ and every $\gamma < \lambda$ there is a γ', $\gamma \le \gamma' < \lambda$, such that $f(\gamma') \in N$.

Example 6 Define $f : \omega \to [0, 1]$ as follows: $f(2n) = \frac{1}{n}$ and $f(2n + 1) = \frac{n}{n+1}$. Then the function f can be extended to an adherent function with domain $\omega + 1$ in two different ways: set $f(\omega)$ to either 0 or 1. However, it cannot be extended to a continuous function with that domain.

The reason 0 (and 1) is an adherence point in the example is that the f-image of every interval $[n, \omega)$ contains the subsequence $a_k = \frac{1}{k}$ (and $b_k = \frac{k}{k+1}$), from some k onwards, and so its closure contains 0 (and 1). On the other hand, f does not converge, since the distances between neighbouring elements converge to 1.

Within infinitary term rewriting, reductions that are adherent but not convergent require proper computational steps at (or near) the root of a term:

Example 7

$$\begin{aligned} H(Z) &\to I(G(Z)) \\ H(G(x)) &\to G(H(x)) \\ G(I(x)) &\to I(G(x)) \\ A(I(x)) &\to A(H(x)) \end{aligned}$$

There is a unique (no choice of redex) non-converging open reduction sequence that starts at the term $A(H(Z))$: this keeps symbols A and Z at the outer points, and grows a block of Gs in the middle. As a result there are infinitely many accumulation points this sequence adheres to, in particular we have $A(H(Z)) \twoheadrightarrow_a A(G^\infty)$, but also $A(H(Z)) \twoheadrightarrow_a A(G^n(I(G^\infty)))$ for any n.

Lemma 6 *A sequence $f : \alpha \to T$ into a metric space T is adherent iff for any limit ordinal $\lambda < \alpha$ the restriction $f|_\lambda$ has a cofinal subsequence with domain ω that converges to $f(\lambda)$.*

Proof If $f|_\lambda$ has a cofinal subsequence that converges to $f(\lambda)$ then that sequence will eventually lie within the f-image of $[\gamma, \lambda)$, and as the sequence converges to $f(\lambda)$, $f(\lambda)$ must lie in the closure of that set.

Conversely, let f be adherent at λ; we can construct a cofinal subsequence $g : \omega \to \lambda$ as follows: let $g(0) = 0$ and then set $g(n+1)$ to be the smallest $k > g(n), k < \lambda$ such that $d(f(k), f(\lambda)) < \frac{1}{n}$. Such k always exist, because otherwise all values in the f-image of $[g(n)+1, \lambda)$ would be distance $\frac{1}{n}$ or more from $f(\lambda)$, so it could not be in their closure. Clearly, $f \circ g$ converges to $f(\lambda)$. □

The argument in the proof is fairly standard and could be generalised from metric spaces to topological spaces that are "first-countable" [11], which means that any point has a countable collection of neighbourhoods such that any neighbourhood of the point has one of them as a sub-neighbourhood. For example, the natural numbers equipped with the cofinite topology (a non-empty set is open iff it is the complement of a finite set) is a topological space that is clearly first-countable but not metrizable.

Given a relation R, the relation $A(R) = \twoheadrightarrow_a$ is defined as follows: $t \twoheadrightarrow_a u$ iff there is an ordinal α and an adherent R-reduction sequence $f : \alpha+1 \to T$ such that $f(0) = t \wedge f(\alpha) = u$. Completely analogous to $\twoheadrightarrow_w$, the function A is monotonic.

Proposition 9 *For any topological space, A is a closure operator.*

Proof Clearly, we can take any $A(R)$ reduction sequence $f : \alpha \to T$ and expand each step $f(\gamma) \twoheadrightarrow_a f(\gamma+1)$ by the R reduction sequence witnessing that step. The resulting sequence g is an R reduction sequence. We need to show that g is adherent at all limit ordinals in its domain.

Limit ordinals in the domain of g can arise in two ways: firstly, they can be of the form $\zeta_\gamma + \lambda$, where ζ_γ is the length of the expanded f-sequence from 0 to γ, and λ is a limit ordinal in the domain of the witnessing sequence g_γ for $f(\gamma) \twoheadrightarrow_a f(\gamma+1)$. In this case g_γ must be adherent at λ and thus so is g at $\zeta_\gamma + \lambda$.

Secondly, μ could be a limit ordinal in the domain of f, and the length of the expanded sequence up to μ is a limit ordinal ζ_μ. Take any neighbourhood N of $f(\mu)$ and any ordinal $\kappa' < \zeta_\mu$. There must be a $\kappa < \mu$ such that $\zeta_\kappa \geq \kappa'$. Since f is adherent at μ there must be a γ, $\kappa \leq \gamma < \mu$ with $f(\gamma) \in N$. But $f(\gamma) = g(\zeta_\gamma)$ and $\lambda > \zeta_\gamma \geq \zeta_\kappa \geq \kappa'$. Hence g is adherent at ζ_μ. □

For metric spaces (more generally, for spaces that are first countable), one has a simpler argument: f is adherent at λ iff a cofinal subsequence of f converges to it—and as cofinal subsequences compose the result follows. However, the proof shows that the property holds for relations on any topological space.

Moreover, on metric spaces A is a closure operator we already encountered: $A = W^\star$. One can observe this as follows:

Proposition 10 *For any topological space T: if $f : \alpha \to T$ is continuous at $x \in \alpha$ then it is also adherent at x.*

Proof The condition (for all U) $x \in \mathrm{Cl}(U) \Rightarrow f(x) \in \mathrm{Cl}(f(U))$ is equivalent to f being continuous at x (folklore). □

Corollary 3 *For any topological space T and any relation R on T, $W^\star(R) \subseteq A(R)$.*

Proof From Proposition 10 it is immediate that any continuous reduction sequence is adherent, hence $W(R) \subseteq A(R)$. Then apply Proposition 9 and monotonicity of W and the Knaster-Tarski Fixpoint Theorem. □

For metric spaces (or indeed first-countable spaces), we can switch the result around:

Proposition 11 *If T is a metric space and R a relation on T, then $W^\star(R) = A(R)$.*

Proof One direction is an instance of Corollary 3.

For the other, we need to show $t\ A(R)\ u \Rightarrow t\ W^\star(R)\ u$; this is shown by induction on α, where α is the smallest ordinal for which there is an adherent sequence $f : \alpha + 1 \to T$ with $f(0) = t$ and $f(\alpha) = u$.

If $\alpha = 0$ then $t = u$ and the result follows by reflexivity of $W^\star(R)$.

If $\alpha = \beta + 1$, i.e. it is a successor ordinal, then $t\ A(R)\ f(\beta)\ R\ u$. By induction hypothesis $t\ W^\star(R)\ f(\beta)$ and any sequence witnessing that can clearly be extended by another R-step.

If α is a limit ordinal then we can use Lemma 6: there must be a cofinal subsequence $g : \omega \to \alpha$ of f that converges to u. This also means $t = f(0) \twoheadrightarrow_a f(g(0)) \twoheadrightarrow_a f(g(1)) \twoheadrightarrow_a \ldots$. The sequences witnessing each $f(g(i)) \twoheadrightarrow_a f(g(i+1))$ can be taken as subsequences of f that are strictly shorter than α, since their lengths are at most $g(n)$ for some n, and $\forall k.\ g(k) < \alpha$. Hence, we can use the induction hypothesis, giving us: $t = f(0)\ W^\star(R)\ f(g(0))\ W^\star(R)\ f(g(1))\ W^\star(R)\ \ldots$. Thus this is a $W^\star(R)$ reduction sequence converging to u and therefore $t\ W^\star(R)\ u$. □

For infinitary term rewriting this means that adherence is indeed what we get when we iterate the weak convergence construction.

When we apply the closure operators we have considered so far $(S,\ W,\ W^\star,\ A)$ to relations we always get pre-orders: all are based on reduction sequences, permit the empty sequence and support the sequential composition of sequences. However, they do not preserve symmetric relations, see the following counterexample.

Example 8 Consider the iTRS (over metric d_∞) with the following two rules:

$$A \to B$$
$$A \to F(B)$$

We consider the symmetric relation $\leftrightarrow_R$, which we get by allowing the rules to be applied forwards as well as backwards. All loop-free reduction sequences of this iTRS are strongly converging and so we have $S(\leftrightarrow_R) = W(\leftrightarrow_R) = A(\leftrightarrow_R)$. The open sequence $f : \omega \rightarrow Ter^{\infty}(\Sigma)$ defined as $f(2 \cdot n) = F^n(B)$ and $f(2 \cdot n + 1) = F^n(A)$ is a reduction sequence for $\leftrightarrow$ and as such it strongly converges to F^{∞}. Hence $B\ S(R)\ F^{\infty}$, etc. As F^{∞} is a normal form w.r.t. $\leftrightarrow_R$ we cannot reverse the relation.

The example shows that the formation of equivalence relations for these operators is a little bit more delicate than the formation of pre-orders. Is the equivalence closure of $S(R)$, $A(R)$, etc. the thing we want, or do we want a notion of equivalence that is closed under these relation-forming operations?

In this section we have looked at several reduction-sequence based closure operators. The list was not meant to be exhaustive—one could conjure up further ones, e.g. use sequences where the indexing function is a *closed map* rather than a continuous one. Instead we will now be looking at closure operators that work more directly on the graph of a relation, i.e. viewing the relation as a set of pairs.

4 Relations

Sequences are not the only thing we can add limits to. The most elementary construction of that ilk is the formation of the topological closure of a set, as long as that set is a subset of a topological space. Regarding relations between topological spaces A and B, that instantly means that we can apply such closures to domain and/or codomain of a relation, as well as the relation as a whole, since $A \times B$ is then also a topological space.

What we are generally looking for are operators X that could take the role of the reflexive transitive closure in the world of infinite terms. That means: $X(R)$ should be a pre-order, permit the approximation of infinite terms by finite terms, and X should be a closure operator. Below we are looking at several candidates that follow this brief:

4.1 Pointwise closure

We can view relations as set-valued functions, and add limits to their range. This leads to the following concept:

Definition 5 A relation R between topological spaces is called *pointwise closed* iff the sets $R^x = \{y \mid x\ R\ y\}$ are all closed.

Note: if a relation R is pointwise closed then the range of Proposition 4 can be extended to arbitrary topological spaces. The reason is this: if an element a occurs repeatedly in a continuous R reduction sequence g without having a last occurrence then a is in every neighbourhood of $g(\lambda)$, where $\lambda = \lim_{g(\kappa)=a} \kappa$. This implies by continuity of g: $g(\lambda) \in \mathrm{Cl}(\{a\})$. If the second occurrence of a in g is at a limit ordinal, then we can replace the value of g at this point by $g(\lambda)$ anyway; otherwise this second occurrence is at a successor ordinal $\phi + 1$ and $g(\phi)\ R\ g(\phi + 1) = a$, and since R is pointwise closed $g(\phi)\ R\ g(\lambda)$. Thus in either case we can replace the second occurrence of a in the sequence by $g(\lambda)$ and remove the other steps leading up to $g(\lambda)$ from g.

Proposition 12 *Being pointwise closed is a closable property of relations.*

Proof Let $A = \bigcap_i R_i$. Then $A^x = \{y \mid x\ A\ y\} = \{y \mid \forall i.\, x\ R_i\ y\} = \bigcap_i R_i^x$. Hence A^x is an intersection of closed sets and therefore closed. □

This allows us to use pointwise closure as a relation-constructing property.

Definition 6 The relation $C(R)$ is defined as the pointwise closure of R. We also define $P(R)$ as $C(R^*)$. Moreover, we write $\twoheadrightarrow_p$ for $P^\star(R)$.

C is clearly a closure operator, P may not be (in general). In particular, P can fail to behave as a closure operator on non-left-linear systems as in this example from [8]:

Example 9 Consider the iTRS with rules:

$$\begin{aligned} A &\to G(A) \\ B &\to G(B) \\ F(x,x) &\to C \end{aligned}$$

Here, the relation $P(\to_R)$ is not transitive, because it relates $F(A,B)$ to $F(G^\infty, G^\infty)$ and $F(G^\infty, G^\infty)$ to C, but not $F(A,B)$ to C.

Proposition 13 *Let R be a relation between topological spaces X and Y. The inverse image operations for R and $C(R)$ are the same on open sets. Hence, R is lsc iff $C(R)$ is lsc.*

Proof Consider a subset $A \subseteq Y$, and let $B_0 = R^{-1}(A)$ and $B_1 = C(R)^{-1}(A)$. As $R \subseteq C(R)$ we have $B_0 \subseteq B_1$. We need to show that if A is open then $B_1 \subseteq B_0$ too.

Suppose there exists $x \in B_1 \backslash B_0$. Then there is a $z \in A$ such that $\neg(x\ R\ z)$ and $x\ C(R)\ z$. Moreover, whenever $x\ R\ y$ then $y \notin A$. This means: $D = \{y \mid x\ R\ y\} \subseteq Y \backslash A$. By definition, if $x\ C(R)\ z$ then $z \in \mathrm{Cl}(D)$. However, as A is open, $Y \backslash A$ is closed and thus: $\mathrm{Cl}(D) \subseteq \mathrm{Cl}(Y \backslash A) = Y \backslash A$ and therefore $z \notin A$—contradiction, there are no such z and x. $\square$

We can explain $t \twoheadrightarrow_p u$ as "t can rewrite to something arbitrarily close to u", but if we want to get any closer we may have to start all over again from t. To put it differently: if $t \to_R^* u_1$, $t \to_R^* u_2, \ldots$ and the u_i form a sequence converging to u then $t \twoheadrightarrow_p u$; moreover, for pattern iTRSs this characterisation is complete (see Theorem 1 below), so that every reduct arises as the limit of reducts that can be reached in finitely many steps.

Proposition 14 *Let T be a metric space and R a relation on T. Then: $A(R) \subseteq P(R)$.*

Proof By induction on the indexing ordinals for the sequences witnessing $t \twoheadrightarrow_a u$. The interesting case for $f : \alpha + 1 \to T$ is when α is a limit ordinal. By induction hypothesis, $f(0) \twoheadrightarrow_p f(\gamma)$, for all $\gamma < \alpha$. By Lemma 6 the restriction of f to α has to contain a subsequence that converges to $f(\alpha)$. But then $f(\alpha)$ has to be in the closure of the $f(\gamma)$ and as $\twoheadrightarrow_p$ is pointwise closed the result follows. $\square$

However, the relations $\twoheadrightarrow_a$ and $\twoheadrightarrow_p$ are not always the same:

Example 10

$$\begin{aligned} A &\to B(A) \\ A &\to C \\ B(C) &\to D(C) \\ B(D(x)) &\to D(D(x)) \end{aligned}$$

In this system we have $A \to_R^* D^n(C)$ for any finite n and therefore $A \twoheadrightarrow_p D^\infty$. But there is no single adherent (or convergent) sequence that can build up to that limit; as soon as a D appears in a reduct of A the reduction sequence is guaranteed to terminate.

For a large class of relations, P is a closure operator:

Theorem 1 *For relations R on any topological space T, if R is lsc, then $P^\star(R) = P(R)$.*

Proof The theorem is clearly equivalent to the transitivity of $P(R)$, as $P(R)$ is reflexive anyway. If $P(R)$ is a pre-order then: $P(P(R)) = C(P(R)^*) = C(P(R)) = C(C(R^*)) = C(R^*) = P(R)$. Hence $P(R)$ is a fixpoint of P.

Since R is lsc so is R^* [16]. Proposition 13 then shows that the pointwise closure also preserves lower semi-continuity.

Consider $a\ P(R)\ b\ P(R)\ c$, and assume $\neg(a\ P(R)\ c)$. Let C be the set $\{d \mid \neg(a\ P(R)\ d)\}$. Using the assumption and the definition of $P(R)$, C must be an open neighbourhood of c. Let $B = P(R)^{-1}(C)$ and $A = P(R)^{-1}(B)$; because $P(R)$ is lsc the sets B and A must be open neighbourhoods of b and a, respectively. By Proposition 13 the inverse images of $P(R)$ and R^* coincide on open sets, which means $a\ R^*\ b'\ R^*\ c'$ for some $c' \in C$, $b' \in B$. By transitivity of R^* we have $a\ R^*\ c'$, and therefore $a\ P(R)\ c'$. Using the definition of C this implies the contradiction $c' \notin C$. □

In [16] a number of conditions are given under which the relation $\rightarrow_R$ is *uniformly lsc*, for a variety of term metrics. For the much weaker condition that $\rightarrow_R$ is lsc it suffices (see Proposition 3) to require that all left-hand sides are patterns, and this result extends to all term metrics.

Corollary 4 *P is a closure operator on pattern iTRSs, for any term metric.*

Proof Consequence of Theorem 1 and Proposition 3. □

Unravelling the proofs of these previous results means that the corollary directly applies for iTRSs with any term metric; but it also applies to iTRSs with other topologies, as long as context application and substitution application satisfy Lemmas 1 and 2 for that topology.

4.2 Topological closure

The pointwise closure adds limits to a relation at the "result side", and stays in this respect still very much within the intuition behind infinitary rewriting. Going beyond that and allowing the input side to change as well leads to fairly unintuitive relations.

For example, one could require that a relation is pointwise closed in either direction. Another closable property on relations between topological spaces A and B is that their set of pairs (their graph) is closed in the product space $A \times B$.

Definition 7 Given a relation R on a topological space A, the relation $T(R)$ is the topological closure (relative to the product space $A \times A$) *of* the reflexive and transitive closure of R. We also write $\twoheadrightarrow_t$ for $T^\star(R)$.

This means: if t_n and u_n are sequences converging to t and u, respectively, and if for all i: $t_i \twoheadrightarrow_t u_i$, then $t \twoheadrightarrow_t u$.

Clearly, closed relations are also pointwise closed and therefore $P(R) \subseteq T(R)$ and also $\twoheadrightarrow_p \subseteq \twoheadrightarrow_t$. Again, the inclusion is proper:

Example 11

$$\begin{aligned} \mathrm{LEQ}(0, x) &\to T \\ \mathrm{LEQ}(S(x), 0) &\to F \\ \mathrm{LEQ}(S(x), S(y)) &\to \mathrm{LEQ}(x, y) \end{aligned}$$

$$
\begin{tikzcd}
& S^\star \arrow[dr, "0"] & & & & & P^\star \arrow[dr, "0"] & \\
S \arrow[ur, "4"] \arrow[dr, "0"'] & & W^\star \arrow[r, "1"] & A \arrow[r, "="] & A^\star \arrow[r, "2"] \arrow[drr, bend right, "0"'] & P \arrow[ur, "3"] \arrow[dr, "0"] & & T^\star \\
& W \arrow[ur, "0"'] & & & & & T \arrow[ur, "0"'] &
\end{tikzcd}
$$

Fig. 1 Hierarchy of closure operators

The infinite term $t = \mathrm{LEQ}(S^\infty, S^\infty)$ only reduces to itself, in a single step, and thus also $\forall u.\, t \twoheadrightarrow_p u \Rightarrow t = u$. We also have $t \twoheadrightarrow_t T$ and $t \twoheadrightarrow_t F$, because the sequences $a_n = LEQ(S^n(0), S^n(0))$ and $b_n = LEQ(S^{n+1}(0), S^n(0))$ both converge to t, but $a_n \twoheadrightarrow_t T$ and $b_n \twoheadrightarrow_t F$.

In contrast to P, the map T is not a closure operator on pattern iTRSs, since $T(\rightarrow_R)$ can much more easily fail to be transitive:

Example 12 Add the rule INF $\rightarrow S$(INF) to Example 11. Writing $\twoheadrightarrow_x$ for $T(\rightarrow_R)$ we have: LEQ(INF, INF) $\twoheadrightarrow_x$ $\mathrm{LEQ}(S^\infty, S^\infty) \twoheadrightarrow_x T$, but we do not have LEQ(INF, INF) $\twoheadrightarrow_x T$.

4.3 Overview

A summary of the relationships between the various closure operators is shown in Fig. 1, where the labels on the arrows mean the following:

- 0 always an inclusion, usually a proper inclusion; this is mostly fairly trivial—for example Corollary 2;
- 1 always an inclusion, an equality for metric spaces (Proposition 11)
- 2 an inclusion for metric spaces (Proposition 14)
- 3 always an inclusion, an equality for pattern iTRS (Theorem 1)
- 4 always an inclusion, an equality for compatible relations (Proposition 8)
- = the equality $A = A^\star$ is Proposition 9

5 Models

The concept of a model allows us to reason semantically about infinite terms. There are different kinds of models one can look at. Perhaps the most important distinction is whether rewrite-steps should be non-trivial at model-level, or whether terms that rewrite to one another should be interpreted as the same value—similar as in denotational semantics of programming languages.

We are looking here at the second case, and thus in particular at Σ-algebras of some kind. The reason for this in the context of this paper is that such models are closely connected to equivalence relations on terms. Each model induces an equivalence relation on terms (the kernel of its interpretation map), but for a class of models we can even construct a closure operator on equivalence relations, which we will consider later with other closure operators.

Σ-algebras alone do not provide an interpretation for infinite terms, unless we equip our class of models (and the interpretation of terms) with some additional structure.

An obvious choice is to require that the models are topological spaces, and that the interpretation map is continuous. It will become clearer later why it makes sense to also consider models with coarser topologies than metric spaces.

The presence of infinite terms complicates the notion of model, firstly because they provide the only non-discrete points in the domain of an interpretation and therefore the only points where an interpretation might fail to be continuous. Secondly, they do not have a direct syntactic presence to which we could latch onto in a definition of model—which we could if, say, all our infinite terms were defined through a fixpoint operator. Thirdly, the syntactic presence they do possess (converging sequences of finite terms) is in its interpretation potentially ambiguous, depending on the model class.

5.1 Preliminaries

In the following we need additional notions from Topology we only use in this section.

Given two topological spaces A and B, the set of continuous functions $A \rightarrowtail B$ can be equipped with the compact-open topology [10]—generated from the subbase that each function set $\{f \mid f(C) \subseteq O\}$ is open for any compact subset C of A and any open subset O of B. Notice that if A is discrete then this coincides with the product topology.

Given a set S, a family $(i \in I)$ of topological spaces Y_i, and a matching family of set-theoretic functions $f_i : S \to Y_i$ the topology over S generated from the subbase of sets of the form $f_i^{-1}(O)$ (with O open in Y_i) is called the *initial topology* for this family.

Initial topologies reflect many separation properties:

Proposition 15 *Let S be equipped with the initial topology of a family of maps $f_i : S \to Y_i, i \in I$. Assume further that whenever $s_1, s_2 \in S$ and $s_1 \neq s_2$ then $f_j(s_1) \neq f_j(s_2)$ for some $j \in I$. Then: the initial topology is T_0, T_1, T_2 or T_3, if all Y_i are, respectively.*

Proof T_0: if $s_1 \neq s_2$ then for some $j \in I$ we have $f_j(s_1) \neq f_j(s_2)$. As Y_j is T_0 there is an open set O in Y_j containing one of the $f_j(s_k)$ but not the other. Consequently $f_j^{-1}(O)$ will be open and contain s_k, but not s_{3-k}.

T_1: the argument is as for T_0 instead that we get two open sets O_1 containing $f_j(s_1)$ and not $f_j(s_2)$, and a dual O_2. The inverse image reflects those properties onto s_1 and s_2.

T_2: the argument is as for T_1 instead that the O_k are disjoint which is again reflected by the inverse image of f_j.

T_3: let $s \in S$ and $C \subset S$ be a closed subset of S not containing s. The argument goes by induction over the formation of closed sets in the initial space. If $C = f_j^{-1}(C')$ for some closed set C' in Y_j then $f_j(s) \notin C'$ and the open neighbourhoods witnessing the separation of C' and $f_j(s)$ in Y_j are mapped by f_j^{-1} to open neighbourhoods witnessing the separation of s and C in S. If $C = C_1 \cup C_2$ then by induction hypothesis s is separable from C_k $(k \in \{1, 2\})$ by open neighbourhoods X_k and E_k. In that case, $X_1 \cap X_2$ and $E_1 \cup E_2$ are the open neighbourhoods separating s from C. If C is an arbitrary intersection of closed sets $C_k, k \in J$ then we must have $s \notin C_n$ for some $n \in J$. By induction hypothesis there are open sets witnessing the separation of s from C_n and these also witness the separation of s from C. □

If B is a bounded metric space with distance function d_B then any function space from A to B can be equipped with the *uniform metric*: $d(f, g) = \sup_{a \in A} d_B(f(a), g(a))$. If A is countable via bijection $h : \omega \to A$, then the infinite product $A \to B$ is also metrizable as $d(f, g) = \sup_{a \in A} \frac{d_B(f(a), g(a))}{h^{-1}(a)}$. The uniform metric is a metrization of the so-called *box topology* for infinite products [11].

For sets of functions equipped with the uniform metric we use the following notations: $A \xrightarrow{u} B$ is the space of set-theoretic function from A to B; if A is a metric space then $A \overset{u}{\hookrightarrow} B$ denotes the space of uniformly continuous functions from A to B, and $A \overset{u}{\rightsquigarrow} B$ the space of non-expansive functions from A to B.

5.2 Equational topological models

As mentioned in the preliminaries, each Σ-algebra comes with an interpretation map $[\![\_]\!]_A : \mathit{Ter}(\Sigma) \to (\mathit{Var} \to A) \to A$ for *finite* terms. As $\mathit{Ter}(\Sigma)$ and Var are discrete, for two of the three "$\to$" in the type of $[\![\_]\!]_A$ it makes no difference if we replaced them by $\rightarrowtail$ (set of continuous functions with compact-open topology); but for the third it does:

Example 13 Consider the signature Σ with unary S. On the interval $[1, 2]$ we define the function g as $g(x) = \frac{x+\frac{2}{x}}{2}$. We define $f(x) = g(x)$, if x is irrational, and $f(x) = g(g(x))$, if x is rational. The Σ-algebra A is set to have carrier set $[1, 2]$ with $S_A = f$.

The interpretation map $[\![\_]\!]_A$ of Example 13 has type $\mathit{Ter}(\Sigma) \to (\mathit{Var} \to A) \to A$, but not type $\mathit{Ter}(\Sigma) \to (\mathit{Var} \to A) \rightarrowtail A$.

Under the first view, there is nothing to show as there are no continuity constraints. Even if we extended the domain of $[\![\_]\!]_A$ to $\mathit{Ter}^\infty(\Sigma)$, and adjusted its type to $\mathit{Ter}^\infty(\Sigma) \rightarrowtail (\mathit{Var} \to A) \to A$ the example would remain valid: the most significant convergent sequence to consider is the sequence $t_n = S^n(x_n)$ with limit S^∞. Although all terms in the sequence use different variables, all interpretations $[\![t_n]\!]_A^\rho$ converge for any ρ to $\sqrt{2}$. But the function f is continuous at exactly one point in its domain, and that is $\sqrt{2}$. So even in the presence of infinite terms that interpretation remains continuous.

The interpretation fails to have type $\mathit{Ter}(\Sigma) \to (\mathit{Var} \to A) \rightarrowtail A$: a sequence v_n of irrational values that converges to 1 gives rise to a converging sequence of valuations ρ_n with $\rho_n(x) = v_n$, with limit $\rho(x) = 1$. But $[\![S(x)]\!]_A^{\rho_n}$ converges to $g(1) = 1.5$ whilst $[\![S(x)]\!]_A^\rho = f(1) = g(g(1)) \approx 1.41666$. Thus, in this view the interpretation fails to be continuous because one of the algebra maps is not - and this applies even in the absence of infinite terms.

Proposition 16 *If A is a topological space and a Σ-algebra and its interpretation function has type $[\![\_]\!]_A : \mathit{Ter}(\Sigma) \to (\mathit{Var} \to A) \rightarrowtail A$ then each function symbol F of Σ is interpreted by a continuous function F_A in A.*

Proof Suppose F has arity n. Let N be a neighbourhood of $F_A(v_1, \ldots, v_n)$. We need to show that there is a neighbourhood M of $(v_1, \ldots, v_n)$ such that $F_A(M) \subseteq N$. As $F_A(v_1, \ldots, v_n) = [\![F(x_1, \ldots, x_n)]\!]_A^\rho$ (for $\rho(x_i) = v_i$) and as $[\![F(x_1, \ldots, x_n)]\!]_A$ is continuous there must be an open neighbourhood M' of ρ such that $[\![F(x_1, \ldots, x_n)]\!]_A\,(M') \subseteq N$. Projecting from M' at the variables $x_1, \ldots, x_n$ gives a set of n-tuples M that contains $(v_1, \ldots, v_n)$. As the projection functions in the product topology are open maps M must be open too and has therefore the required properties. □

To provide a full interpretation of finite and infinite ground and non-ground terms, we can enforce an interpretation map by viewing it as a primitive rather than a derived notion:

Definition 8 Given a metric signature Σ_m, a Σ_m*-algebra* is given by a topological space A and a function $[\![\_]\!]_A : \mathit{Ter}^m(\Sigma) \rightarrowtail (\mathit{Var} \to A) \rightarrowtail A$, (notation: $[\![t]\!]_A^\rho$ for $[\![\_]\!]_A(t)(\rho)$) such that two conditions hold:

1. $\{\![x]\!\}_A^\rho = \rho(x)$, for all $x \in \mathit{Var}$ and all $\rho : \mathit{Var} \to A$;
2. $\{\![\theta(t)]\!\}_A^\rho = \{\![t]\!\}_A^\varrho$, where $\forall x \in \mathit{Var}.\ \varrho(x) = \{\![\theta(x)]\!\}_A^\rho$, for all terms $t \in \mathit{Ter}^m(\Sigma)$, all substitutions $\theta : \mathit{Var} \to \mathit{Ter}^m(\Sigma)$, and all $\rho : \mathit{Var} \to A$.

Homomorphisms between Σ_m-algebras A and B are continuous functions $h : A \rightarrowtail B$ such that for all $t \in \mathit{Ter}^m(\Sigma)$ and $\rho : \mathit{Var} \to A$ we have:

$$h\left(\{\![t]\!\}_A^\rho\right) = \{\![t]\!\}_B^{h\circ\rho}$$

Note that this is essentially an instance of the notion of Eilenberg-Moore-algebra of a monad in category theory [27, p. 136]. The monad in this case are infinite terms over an arbitrary set X (generalised variables); the unit of the monad is the embedding of X as variables into the term set, the monad's multiplication (or join) is substitution; in [26] Lüth considered the corresponding monad for finite terms. The product topology for $(\mathit{Var} \to A) \to A$ inherits many topological properties from A: if A is T_0, T_1, T_2 or T_3 then so is $(\mathit{Var} \to A) \to A$ [11].

Our notion of Σ_m-algebra is compatible with the notion of a Σ-algebra:

Proposition 17 *Every Σ_m-algebra A gives rise to a Σ-algebra B such that $[\![\_]\!]_B$ is the restriction of $\{\![\_]\!\}_A$ to $\mathit{Ter}(\Sigma)$.*

Proof For any function symbol G or arity n we can define $G_B(v_1, \ldots, v_n)$ as $\{\![G(x_1, \ldots, x_n)]\!\}_A^\rho$ with $\rho(x_i) = v_i$. The coherence conditions for Σ_m-algebras give us (for any valuation ρ) the following:

$$\{\![G(t_1, \ldots, t_n)]\!\}_A^\rho = \{\![G(x_1, \ldots, x_n)]\!\}_A^\pi = G_B(\{\![t_1]\!\}_A^\rho, \ldots, \{\![t_n]\!\}_A^\rho)$$

where $\pi(x_i) = \{\![t_i]\!\}_A^\rho$. □

For an arbitrary topological space it does not generally suffice to provide an interpretation for finite terms, and hope the interpretation of infinite terms derives through topological magic. Convergence is generally too weak a concept to deliver that. However, for Hausdorff spaces we have this:

Proposition 18 *Let A be a Σ-algebra with a Hausdorff topology. If the interpretation of finite terms $[\![\_]\!]_A$ can be extended to a Σ_m-algebra map then this extension is unique.*

Proof Suppose we had two such continuous extensions f and g, i.e. they coincide on finite terms. Take any infinite term t, we need to show: $f(t) = g(t)$. As any term, t can be approximated by a sequence of finite terms t_n. We have $f(t_n) = g(t_n)$, pointwise. By continuity of f and g, the sequence $f(t_n)$ converges in $(\mathit{Var} \to A) \rightarrowtail A$ to both $f(t)$ and $g(t)$, and as the topology of $(\mathit{Var} \to A) \rightarrowtail A$ inherits Hausdorff-status from A we have $f(t) = g(t)$. □

The uniqueness of the extension does not mean that it always exists, and one has to be careful that the continuity of the extended interpretation map extends to the model as a whole:

Example 14 Consider the signature Σ with unary S, nullary 0, and term metric $S_m(x) = x/2$. We define a Σ-algebra A on the interval $[0, 1]$ with its standard topology, with $0_A = 0$ and $S_A(x) = x^2$. Although 0 is a fixpoint of S_A, A cannot be extended to a Σ_m-algebra because S_A has another fixpoint with 1. Consider the sequence of terms $S^n(x)$ (converging to S^∞) with the valuations $k0(x) = 0$ and $k1(x) = 1$; $\left[\!\left[S^n(x)\right]\!\right]_A^{k0} = 0$ and $\left[\!\left[S^n(x)\right]\!\right]_A^{k1} = 1$. Thus, whatever $\{\![S^\infty]\!\}_A^{k0} = \{\![S^\infty]\!\}_A^{k1}$ is, a sequence that is constantly 0 or constantly 1 would have to converge to both, which is impossible in a T_1 space.

We can turn the example into a Σ_m-algebra by limiting the carrier set to $[0, 1)$. Valuation $k1$ would no longer exist in this model. and sequences of valuations in $\mathit{Var} \to [0, 1)$ that converge to $k1$ in $\mathit{Var} \to [0, 1]$ diverge in $\mathit{Var} \to [0, 1)$.

Definition 9 Any Σ_m-algebra A (non-empty class of Σ_m-algebras B) induces an equivalence $\sim_A$ ($\sim_B$) on $Ter^m(\Sigma)$ where $t \sim_A u \iff [\![t]\!]_A = [\![u]\!]_A$; $\sim_B$ is the intersection of all $\sim_X$ for $X \in B$.

Notice that this definition is completely analogous to the standard notion of a congruence relation induced by a class of Σ-algebras in Universal Algebra [28, Definition 5.1.8] or, more generally in Model Theory, to the concept of a theory derived from a class of models [30, Definition 2.3.1].

Proposition 19 *Relation $\sim_A$ (and similarly $\sim_B$) is a strong congruence relation in the sense: for all $t, u \in Ter^m(\Sigma)$ such that $t \sim_A u$ and all substitutions θ, υ : such that $\forall x \in Var.\, \theta(x) \sim_A \upsilon(x)$ we have $\theta(t) \sim_A \upsilon(u)$.*

Proof The coherence condition for Σ_m-algebras give us: $[\![\theta(t)]\!]_A^\rho = [\![t]\!]_A^\psi$, where for all $x \in Var$ we have $\psi(x) = [\![\theta(x)]\!]_A^\rho$, and similarly, $[\![\upsilon(u)]\!]_A^\rho = [\![u]\!]_A^\phi$, where $\phi(x) = [\![\upsilon(x)]\!]_A^\rho$. As $\theta(x) \sim_A \upsilon(x)$ by assumption we have $[\![\upsilon(x)]\!]_A = [\![\theta(x)]\!]_A$, and thus applying the functions to ρ gives the same result. Thus ϕ and ψ are pointwise the same, and as $[\![t]\!]_A = [\![u]\!]_A$ the result follows. Clearly, strong congruences are preserved by arbitrary intersections, so the result for $\sim_B$ follows. □

If we are specifically interested in models that as topological spaces are T_0, T_1, etc., then we can use initial topologies to construct initial models, i.e. initial Σ_m-algebras in the class of Σ_m-algebras that satisfy a set of equations.

Proposition 20 *Any set of $Ter^m(\Sigma)$-equations E has an initial T_i model, for $i \in \{0, 1, 2, 3\}$.*

Proof By construction of the model. Let M be the class of all Σ_m-algebras that are T_i spaces and which satisfy E, i.e.: for all equations $t = u$ and all $A \in M$ and all valuations $\rho : Var \to A$ we have $[\![t]\!]_A^\rho = [\![u]\!]_A^\rho$. From this we can form the quotient algebra $Ter^m(\Sigma)/{\sim_M}$—clearly, all interpretations into models in M factor through that quotient algebra.

From here Proposition 15 applies, because every two distinct $\sim_M$-equivalence classes will be distinguished in at least one model. □

There are two main reasons why we want to consider models with weaker topologies than metric spaces:

- There is a potential semantic gap between transfinite rewriting and models [18], in the sense that the induced equivalence of a class of models often equates more terms than the equivalence relations we build on top of transfinite rewrite relations; widening the class of models narrows that gap.
- An important special case of the previous issue is when $\sim_M$ equates all terms, so we really had inconsistent requirements. For example, the presence of a collapsing rule w.r.t. metric d_∞ makes this happen in the class of models that are Hausdorff spaces. But the inconsistency is not global, it is relative to the class of models under consideration.

5.3 Equational metric models

As metric spaces are special cases of topological spaces we can simply instantiate the notion of topological model (for a set of equations) with metric spaces. However, the constructions of Proposition 20 would not carry over to metric spaces, because the initial topology derived from all metric models may not be metrizable. This is so because continuous functions do not reflect enough structure to be able to construct a metric from that.

Example 15 Consider the first uncountable ordinal ω_1. We can define maps f_α from it to countable ordinals $\alpha + 1$ by setting $f_\alpha(x) = x$ if $x < \alpha$ and $f_\alpha(x) = \alpha$ if $x \geq \alpha$. The order topology of each $\alpha + 1$ is metrizable by Urysohn's metrization theorem [35], the initial topology induced on ω_1 by the family f_α, $\alpha < \omega_1$, is its order topology which is not metrizable.

To still get "initial models" for metric spaces one has to require that semantic interpretations preserve additional structure.

One advantage of metric models over topological ones is that the interpretation of infinite terms can be *constructed*, rather than merely requiring that they have one. This construction goes via metric completion, a process that also requires that semantic interpretations preserve more structure than continuous functions do.

Before considering these issues in more detail, we reflect on the literature of the subject as semantic interpretations of infinitary term rewriting systems have been largely confined to metric models. The main differences between these approaches have to do with the extra structure that is employed to establish interpretations for infinite terms. A variety of ideas have been tried:

In [7] the issue of interpreting infinite terms was side-stepped: only interpretations of finite terms were provided a priori (thus rules were not allowed infinite terms in their right-hand sides). In fact, in the special case of equational models (rather than partially ordered ones), the construction in [7] specialises exactly to ordinary Σ-algebras satisfying a set of equations. As a consequence, infinite terms may not just have no interpretation to begin with, it may be impossible to extend the interpretation of finite terms to infinite ones in a continuous manner, and the semantic interpretation of converging rewrite sequences may not converge.

Zantema interpreted iTRSs in weakly monotone Σ-algebras with some extra structure [36]. A special case are ordinary Σ-algebras since the ordering can be chosen to be equality. The extra structure required on a model A in that paper consists of the following: (i) a metric d_A; (ii) continuity of F_A for every function symbol F, w.r.t. the topology induced by the metric; (iii) the interpretation of the sequences $\mathrm{trunc}(t, n)$ in A converges for every infinite ground term t. Here, $\mathrm{trunc}(t, n)$ replaces all subterms of t at depth n with the fixed constant c.

Although condition (iii) gives us an interpretation of infinite ground terms this interpretation may not be continuous, because other sequences of finite ground terms converging to the same infinite term may fail to converge in A; thus, Endrullis et al. [9] modified condition (iii) by requiring instead that *all* converging sequences of ground terms converge in A.

By sticking in the conditions to ground terms, aspects of A that are not in the image of the interpretation are ignored. In particular, continuity of the interpretation function is guaranteed for *substitutions* [9, Lemma 3.3], but not for *valuations*:

Example 16 Consider the signature with nullary 0 and unary S. We can build a metric model A for this (in the sense of [9]) by using the field Z_7 (integers modulo 7) as carrier set, setting $0_A = 0$ and $S_A(x) = x^2 \pmod 7$, and equipping it with the discrete metric. This gives an interpretation of $S_A^\infty = 0$. The sequence of open terms $t_0 = x,\ t_{i+1} = S(t_i)$ (which has S^∞ as a limit) interprets in the model as a sequence of continuous functions in $A \to A$, but this sequence is not converging to the constant function $x \mapsto 0$ as we might expect. Instead, it is not converging at all, as it eventually flip-flops between x^2 and x^4, since $(x^2)^2 = x^4$ and $(x^4)^2 = x^8 = x^2 \pmod 7$ by Fermat's "little theorem". Moreover, the model has with 1 another fixpoint of S_A and thus a non-canonical interpretation of S^∞ too.

An alternative to requiring *a priori* that infinite terms have an interpretation is to derive such an interpretation through metric completion, which is the approach taken in [18].

For this to work, the interpretation function would have to be uniformly continuous. In [18] this was achieved by requiring a connection between distances in the model and the term metric, which would ensure that the interpretation function was non-expansive [18, Lemma 6.2]. However, one can make the definition of a semantic interpretation more lightweight by requiring non-expansiveness up-front instead of including a particular means to achieving that in the definition. As a consequence, the notion of *metric model* from [18] is an instance of the notion of *metric algebra* below.

Definition 10 Given a metric signature Σ_m, a *metric algebra* is a Σ-algebra A, such that A is a bounded metric space, and the interpretation map $[\![\_]\!]_A$ has type $Ter(\Sigma) \overset{u}{\rightsquigarrow} (Var \to A) \overset{u}{\hookrightarrow} A$.

To explain the choice of function spaces and topologies in the definition: requiring the interpretation map $[\![\_]\!]_A$ to be non-expansive means here that that $d_m(t,u) \geq d_A([\![t]\!]^\rho_A, [\![u]\!]^\rho_A)$, for any valuation ρ, where d_A is the distance function on A. That in itself already entails that A is bounded, because if $\rho(x) = v$, $\rho(y) = w$ then $d_A(v,w) = d_A([\![x]\!]^\rho_A, [\![y]\!]^\rho_A) \leq d_m(x,y) = 1$.

Choosing $\overset{u}{\hookrightarrow}$ for the second half of the interpretation means that the interpretation of terms is uniformly continuous, pointwise—so each finite term t gives rise to a uniformly continuous map $[\![t]\!]_A$ which maps the interpretation of t against converging sequences of valuations to the interpretation of t against their limit. In particular, Proposition 16 applies—all function symbols are interpreted by continuous functions. The uniformity allows us to extend the interpretation in an automatic way.

For the definition of metric algebra it makes no real difference whether we use standard $Var \to A$ or box topology $Var \overset{u}{\to} A$ for infinite products. The reason is that $Ter(\Sigma)$ only contains finite terms and these have only finitely many variables. So, for any $t \in Ter(\Sigma)$ and any sequence of valuations ρ_n that is converging in $Var \to A$ (to ρ) but not in $Var \overset{u}{\to} A$, there is another sequence of valuations π_n that is uniformly converging to ρ such that $[\![t]\!]^{\rho_n}_A = [\![t]\!]^{\pi_n}_A$, pointwise. For similar reasons, $[\![t]\!]_A$ is uniformly continuous w.r.t. the metrization of the product topology iff it is w.r.t. the uniform metric. The choice of product metric *would* make a difference though if we made the stronger requirement that $[\![t]\!]_A$ be non-expansive; this could be made to work with the uniform metric but not with standard products, because that would only permit singleton spaces: if valuations ρ and ϱ only differed at variable x_n then $d(\rho,\varrho) = \frac{1}{n} \cdot d_A(\rho(x_n), \varrho(x_n))$, the non-expansiveness condition for the interpretation of x_n would give us $d_A(\rho(x_n), \varrho(x_n)) \geq n \cdot d_A(\rho(x_n), \varrho(x_n))$ and thus $d_A(\rho(x_n), \varrho(x_n)) = 0$.

Proposition 21 *If A is a metric algebra for Σ_m then its metric completion $A^\bullet$ is a Σ_m-algebra.*

Proof Because $[\![\_]\!]_A$ is non-expansive it is also uniformly continuous and so it can be lifted (uniquely) to the respective metric completions of $Ter(\Sigma)$ and $(Var \to A) \overset{u}{\hookrightarrow} A$—and this lifting preserves the non-expansiveness of the map. The metric completion of $Ter(\Sigma)$ w.r.t. d_m is $Ter^m(\Sigma)$. The metric completion lifts $(Var \to A) \overset{u}{\hookrightarrow} A$ to $(Var \to A) \overset{u}{\hookrightarrow} A^\bullet$ where $A^\bullet$ is the metric completion of A [16, Proposition 1]. As every map in $(Var \to A) \overset{u}{\hookrightarrow} A^\bullet$ is uniformly continuous it can itself be uniquely lifted to the metric completions of $Var \to A$ and $A^\bullet$ which overall gives us the type $Ter^m(\Sigma) \overset{u}{\rightsquigarrow} (Var \to A^\bullet) \overset{u}{\hookrightarrow} A^\bullet$.

The coherence law $[\![\theta(t)]\!]^\rho_A = [\![t]\!]^{\rho\circ\theta}_A$ holds because of uniform continuity: if t_i is a sequence of finite terms converging to t, and θ_i a sequence of pointwise finite substitutions converging to θ then for any $\epsilon > 0$ and for sufficiently large n $d_A([\![\theta(t)]\!]^\rho_A, [\![\theta_n(t_n)]\!]^\rho_A) < \epsilon$

and $d_A(\llbracket t\rrbracket_A^{\rho\circ\theta}, \llbracket(t_n)\rrbracket_A^{\rho\circ\theta_n}) < \epsilon$. As the coherence law holds for finite terms and substitutions this implies $d_A(\llbracket\theta(t)\rrbracket_A^{\rho}, \llbracket t\rrbracket_A^{\rho\circ\theta}) < 2\cdot\epsilon$. As ϵ is arbitrary the two interpretations have to be the same. □

By continuity of the interpretation, if a metric algebra satisfies an equation (or set of equations) then so does its metric completion.

For Proposition 21 to hold we would only need that $\llbracket\_\rrbracket_A$ is uniformly continuous. The stronger requirement that the interpretation is non-expansive gives us something extra when we interpret equations.

Definition 11 Let $E \subseteq Ter^m(\Sigma) \times Ter^m(\Sigma)$ be a set of equations. A *metric model of* E is a metric algebra A (for Σ_m) such that for all $(t, u) \in E$ and all valuations $\rho : Var \to A$ we have $\llbracket t\rrbracket_{A^\bullet}^{\rho} = \llbracket u\rrbracket_{A^\bullet}^{\rho}$, where $A^\bullet$ is the metric completion of A.

Theorem 2 *Any set of* $Ter^m(\Sigma)$*-equations* E *has an initial metric model.*

Proof The proof goes by construction of the model I. The carrier set of I is $Ter^m(\Sigma)/{\sim_M}$ where M is the class of all metric models of E.

Since $E \subseteq \sim_M$ we have that I satisfies E under $[\_]_{\sim_M}$ as its interpretation map.

We equip I with a metric d_I as follows:

$$d_I([t]_{\sim_M}, [u]_{\sim_M}) = \sup_A d'_A(\llbracket t\rrbracket_A, \llbracket u\rrbracket_A),$$

where A ranges over all metric models of E, and d'_A is the uniform metric for the function space $(Var \to A) \to A$ derived from the metric d_A on A.

To check that d_I is indeed a metric; the triangle inequality: $\forall A.\ d_A(t, u) \leq d_A(s, t) + d_A(s, u)$ implies $\forall A.d_A(t, u) \leq d_I(s, t) + d_I(s, u)$ and thus also $d_I(t, u) = \sup_A(d_A(t, u)) \leq d_I(s, t) + d_I(s, u)$.

Checking the zero-axiom: $d_I(t, u) = 0$ if $d_A(t, u) = 0$ for all A, i.e. if $t = u$ in all models. Hence $t = u$ in I too. If $t = u$ in I then $t = u$ in all A, hence $d_A(t, u) = 0$ in all A and so $d_I(t, u) = 0$.

Because each $\llbracket\_\rrbracket_A$ map is non-expansive, we have $d'_A(\llbracket t\rrbracket_A, \llbracket u\rrbracket_A) \leq d_m(t, u)$ for all A and therefore $d_I([t]_{\sim_M}, [u]_{\sim_M}) \leq d_m(t, u)$ which means that quotienting by the equivalence is also non-expansive.

Clearly, the interpretation map $\llbracket\_\rrbracket_A$ of any metric model factors uniquely through I as $h_A \circ [\_]_{\sim_M}$ where h_A is also non-expansive. □

Theorem 2 is the reason why metric models are defined here with non-expansive interpretations rather than uniformly continuous ones. If we used the construction of the proof w.r.t. uniformly continuous interpretations then the metric of the quotient would be discrete, and then the quotienting map would not be continuous.

Because we are measuring satisfiability only w.r.t. the metric completion the construction does not distinguish between a model and its metric completion.

6 Notions of equivalence

In the presence of infinite terms, ordinary congruence relations fail to capture what is needed for equational reasoning in infinitary rewriting as equivalence closure is an inductive concept, not a coinductive one. This problem shows up in two separate ways: (i) for including transfinite reductions in the equivalence, and (ii) for allowing infinitely many subterm changes in a term of infinite size—and combinations of those two.

Suppose we have a compatible equivalence relation $\sim$ and two infinite terms of the forms $t = C[t_1, t_2, \ldots]$ and $u = C[u_1, u_2, \ldots]$ where the t_i are $\sim$-equivalent to the u_i. We do not automatically have that t and u are $\sim$-equivalent. If the equivalence $\sim$ is $E(X(R))$ and $X(R)$ is confluent (for $X \in \{S, W, A, P, T\}$) then, for all term metrics, there are common $X(R)$ reducts $s_1, s_2, \ldots$ for the t_i and u_i, and by compatibility $t\ \ X(R)\ \ C[s_1, s_2, \ldots]\ \ X(R)^{-1}\ \ u$, and thus t and u are equivalent. However, when $X(R)$ is not confluent then there is a potential problem: the equivalence proofs in $E(X(R))$ may have unbounded length. Thus, although a single $X(R)$ step can advance all proofs of $t_i\ \ E(X(R))\ \ u_i$ by one step in parallel this is insufficient to reach the limit.

Example 17 Extend Example 8 with a binary symbol G. We have $A \leftrightarrow_R^* F^n(B)$ for any n. The only transfinite reductions in this system are parallel one-step reductions, so $\twoheadrightarrow_s = \twoheadrightarrow_t$. We can define the following infinite terms: $t = G(A, t)$ and $u_k = G(F^k(B), u_{k+1})$ and $v_k = G(x_k, v_{k+1})$. Then there are substitutions θ and υ such that $t = \theta(v_0)$ and $u_0 = \upsilon(v_0)$ and for all variables x_k: $\theta(x_k) \leftrightarrow_R^* \upsilon(x_k)$. But t and u_0 are not related by the equivalence closure of $\twoheadrightarrow_s$.

A further consequence of Example 17 is that the equivalence $E(\twoheadrightarrow_s)$ differs from the equivalence $\sim_A$ induced from its quotient model $Ter^\infty(\Sigma)/E(\twoheadrightarrow_s)$.

Any equivalence relation $\sim$ on a topological space A induces a canonical topology on the quotient $A/\sim$: a set of equivalence classes is open iff their union is open in the topology of A. This establishes the finest topology that makes the projection map $[\_]_\sim : A \to A/\sim$ continuous [11, Theorem 13.1]. The continuity of the projection map means that if we have *any* converging sequence $f(n)$ in A then $[f(n)]_\sim$ is converging in $A/\sim$.

Example 18 Consider the iTRS with the single rule $C \to S(C)$. Take as equivalence $\approx$ the congruence closure of the equation $C = S(C)$. The reduction sequence $C \to S(C) \to S(S(C)) \to \ldots$ converges to S^∞. By continuity, the sequence $[C]_\approx, [S(C)]_\approx, [S(S(C))]_\approx, \ldots$ converges to $[S^\infty]_\approx$. However, $[C]_\approx = [S(C)]_\approx = [S(S(C))]_\approx, \ldots$ and $[C]_\approx \neq [S^\infty]_\approx$.

Example 18 shows that quotient spaces can have very poor separation properties, e.g. in the example $Ter^\infty(\Sigma)/\approx$ is not T_1. These separation properties closely correspond to properties of equivalence relations, in the sense that they indirectly provide recipes for adding limits to an equivalence.

6.1 Closure operators for equivalences

Given any relation R on any set, the equivalence closure of R is $E(R) = (R \cup R^{-1})^*$. In infinitary term rewriting the relation $E(S(R))$ is often used implicitly as an equivalence, because so-called transfinite confluence [20] is just the (ordinary) confluence of the transfinite rewrite relation. What is odd about this construction is that both E and S use a form of many-step reasoning, but not the same one: S uses strongly converging sequences, E finite sequences. For finitary rewriting there is no such discrepancy, and $E(R^*) = E(R)$.

For infinitary rewriting we are looking for something similar, a single principle that works for both. Given any monotonic and increasing function on relations F we can define such an operator F_E as follows: $F_E(R) = (F \circ E)^\star(R)$. In this construction it suffices if F is only defined on equivalence relations, as we only ever apply it to the result of E. Thus $F_E(R)$ will be the smallest equivalence relation containing R that is also a fixpoint of F.

As it may take an infinite number of applications of $F \circ E$ to reach a fixpoint, it is of some interest to find simpler constructions of that fixpoint. In [18] the author gives some conditions under which $F_E(R) = E(F(R))$, for various F. In particular if $F(R)$ is confluent and usc then this holds for $F \in \{P, T^\star\}$. Unfortunately, there are no known techniques to establish the usc property for any of the relation constructions considered here. One can show that the reflexive closure of $\rightarrow_R$ is usc (for any term metric), provided the number of rules is finite, but usc relations are generally not preserved by transitive closure. For Example 10, $\rightarrow_R^*$ is not usc and none of our constructions that "add limits to the right-hand side" are usc; $T(\rightarrow_R) = T^\star(\rightarrow_R)$ is usc for the example.

6.2 Using the closure principles from rewriting

So far, we have encountered various monotonic functions on relations that "added limits" to a relation. The diversity of equivalence relations constructed this way is not quite as wide as that of rewrite relations:

Proposition 22 *On metric spaces, the equivalence relation formations W_E, A_E and P_E coincide.*

Proof The equation $W_E(R) = A_E(R)$ is a consequence of Proposition 11. To prove $W_E(R) = P_E(R)$ we only need to show that the equivalence classes of $W_E(R)$ are closed. Suppose b is in the closure of an $W_E(R)$ equivalence class B. Then B must contain an element from every neighbourhood of b, in particular for each $n \in \omega$ there must a $b_n \in B$ such that $d(b_n, b) < 2^{-n}$. For the sequence b_n we have $b_j\ W_E(R)\ b_{j+1}$ as all b_j inhabit the same equivalence class. Since $W_E(R)$ is closed under W and the sequence b_n converges to b it follows that $b_0\ W_E(R)\ b$ and therefore $b \in B$. $\square$

Again, the result could be generalised to first-countable spaces.

While these notions of equivalence coincide, the equivalence $S_E(R)$ is often weaker. Typically it fails to relate the limit of a weakly (but not strongly) convergent sequence to its approximants, see for instance Example 3 for which $F(A, x)\ S_E(R)\ F(A, D^\infty)$ does not hold.

It is not trivial to come up with examples in which T_E and P_E differ. Here is one.

Example 19 Suppose an iTRS has rules $F^n(A) \rightarrow G^n(A)$ for every $n \in \omega$. Then $F^\infty\ T_E(\rightarrow_R)\ G^\infty$ but these two terms remain unrelated by $P_E(\rightarrow_R)$.

6.3 Equivalences through semantic interpretation

Instead of directly operating with equivalence relations syntactically, we could require that the quotient space $T/\!\sim$ have certain topological separation properties, or that at least it could be canonically be supplied with those.

The idea is to derive such an equivalence formation from a class of models: if terms t and u are interpreted the same in all models of the class then we should regard them the same, so we might as well include their equivalence. If in addition all those models interpret a certain equivalence as equality then the derived equivalence includes the original equivalence. Thus, any class of models also defines a closure operator on equivalence relations:

Definition 12 Let $\mathcal{M}$ be a class of Σ_m-algebras. Given an equivalence $\sim$ on $Ter^m(\Sigma)$, we write:

$$A \models \sim \iff \forall t, u \in Ter^m(\Sigma).\, t \sim u \Rightarrow [\![t]\!]_A = [\![u]\!]_A$$

The equivalence $\mathcal{M}(\sim)$ is then defined as follows:

$$\mathcal{M}(\sim) = \sim_B \qquad \text{where} \quad B = \{X \in \mathcal{M} \mid X \models \sim\}$$

$$\begin{tikzcd}
\Sigma_E^m \arrow[r] & 0_E \arrow[r] & 1_E = P_E \arrow[r] & T_E \arrow[r] & 2_E \arrow[r] & 3_E \\
 & S_E \arrow[ur] & & & &
\end{tikzcd}$$

Fig. 2 Hierarchy of closure operators for equivalences

We write Σ_E^m for the case that the class $\mathcal{M}$ contains all Σ_m-algebras.

If we restrict the class of models, by their topological separation properties, we get more expressive closure operators. This gives rise to closure operators $0_E, 1_E, 2_E, 3_E$, when the class of models is restricted to $T_0,\ T_1,\ T_2$ or T_3 spaces, respectively. These operators merge equivalence classes whenever the separation principles available in those classes cannot separate them.

Comparing these with the closure operators we have already seen gives us the hierarchy in Fig. 2 (see also Sect. 5 in [18]), where the arrows mean "inclusion". Most of these inclusions follow immediately from the corresponding inclusions of their associated model classes. The equality $1_E = P_E$ is well-known [11, p107], and fairly obvious: a quotient space of an equivalence $\sim$ is T_1 iff all its equivalence classes are closed, which is the case iff $\sim$ is pointwise closed. The embedding from S_E to P_E is a consequence of the corresponding embedding from S to P, and the monotonicity of both E and the fixpoint construction.

The subtle distinctions at the upper end of the hierarchy appear to have little practical relevance for iTRSs. For example, if one equivalence class of $\sim$ contained all terms of the form $F(x_n, G^n(x))$ and another accordingly all terms of the form $F(x_n, G^n(y))$ then 2_E would have to merge these classes into one, whilst T_E may not.

At the lower end, the distinctions are more significant. In particular, assuming that $\Sigma_E^m(\sim) = \sim,\ 0_E$ can also be defined as:

$$t\ 0_E(\sim)\ u \iff t \in \mathrm{Cl}([u]_\sim) \wedge u \in \mathrm{Cl}([t]_\sim)$$

The reason is that the condition on the right is topological indistinguishability of $[u]_\sim$ and $[t]_\sim$ in the quotient space, which is an equivalence relation; quotienting by that equivalence always gives a T_0 space.

Although 0_E is a fairly weak extension principle for equivalences, usually weaker than S_E, there are situations where the two are incomparable:

Example 20 Consider the following iTRS (for metric d_∞):

$$\begin{aligned}
F(x) &\to A \\
F(F(x)) &\to F(x) \\
C &\to G^\infty \\
A &\to S(A) \\
A &\to B \\
S(B) &\to G(B) \\
S(G(x)) &\to G(S(x))
\end{aligned}$$

We have, for any $n,\ F^\infty \to_R A \to_R^* S^n(A) \to_R^* G^n(B)$, and also $F^n(C) \to_R^* C \to_R G^\infty$. Thus the equivalence class of F^∞ w.r.t. $E(\to_R)$ contains all $G^n(B)$, and therefore all neighbourhoods of the class of G^∞ will contain that class, and the reverse is also true. Therefore, F^∞ and G^∞ are related by $0_E(\to_R)$, but they are not related by $S_E(\to_R)$.

On the other hand, $A \twoheadrightarrow_s S^\infty$ and thus the two terms are also related by $S_E(\to_R)$, but they are not related by $0_E(\to_R)$.

6.4 Consistency

In finitary term rewriting, the derivability of $x =_R y$ is used as the standard criterion for inconsistency. This form of inconsistency does not mean that a specification has no models, merely that all its models are trivial (with singleton sets as carrier sets). Moreover, if all models of $\sim$ in $\mathcal{M}$ are trivial then $\mathcal{M}(\sim)$ will relate x to y.

For infinitary rewriting over metric d_∞ consistency becomes trivial for T_1 models: the relation $1_E(\to_R)$ (call it $\sim_1$ for brevity here) contains the pair (x, y) iff R contains a collapsing rule. The reason: if it does contain a collapsing rule $C[x] \to x$ then $x \sim_1 C[x] \sim_1 C[C[x]] \sim_1 \cdots \sim_1 C^\infty$ and thus $x \sim_1 C^\infty \sim_1 y$, if it does not then each set $\{x\}$ remains an equivalence class of its own, since x is a discrete point in $Ter^\infty(\Sigma)$: thus $\{x\}$ is closed and no set not containing x has it in its closure.

This is not only a feature of 1_E: for metric d_∞, for all closure operators in the hierarchy from S_E to 3_E consistency only hinges on the presence/absence of collapsing rules.

For other term metrics, or even non-metric topologies on terms, the situation is a little bit more delicate, because not all collapsing rules are problematic and neither are all hypercollapsing terms: take any consistent finite TRS and equip it with the discrete metric; the metric completion will create no infinite terms and the system remains consistent (with $E(\to) = S_E(\to) = 3_E(\to) = \ldots$), even it has hypercollapsing terms.

Definition 13 Let $t \in Ter^m(\Sigma)$. We assume some topology on $Ter^m(\Sigma)$, not necessarily its metric topology. t is called a *collapsing tower* for a relation R if we have:

- a function $f : \omega \to Pos(t)$, such that
- $\forall n \in \omega.\, \exists w_n \neq \langle\rangle.\, f(n) \cdot w_n = f(n+1)$, and
- $f(0) = \langle\rangle$, and
- for each open neighbourhood A of t there is an $n \in \omega$ such that $t[\;]_{f(n)}$ is safe for A
- for each $u \in Ter^m(\Sigma)$ and each $n \in \omega$ we have $t|_{f(n)}[u]_{w_n}\ R\ u$.

The first condition on f ensures that it is an *infinite path* in the sense of [19], the last that it is made out of the composition of collapsing contexts. The condition on neighbourhoods ensures that if we place arbitrary terms on positions arbitrarily deep along that infinite path we get a sequence that (strongly) converges to t. For term metrics that is automatic. We would also get this for the Scott-topology for partial terms: an open set containing t will also contain $t[u]_p$ for all u and all but finitely many p; we would not get this property for the Alexandrov-topology as this is discrete on $\perp$-free terms.

Clearly, in $Ter^\infty(\Sigma)$ under its standard topology we can build collapsing towers iff we have collapsing rules. For other term metrics this is not automatic, because their metric completion can fail to generate those towers. For example, if we have the rule $F(x) \to x$ with term metric $F_m(\epsilon) = \epsilon$ then F^∞ is not in $Ter^m(\Sigma)$.

Proposition 23 *Let R be a rewrite relation on $Ter^m(\Sigma)$, let $t \in Ter^m(\Sigma)$ be a collapsing tower w.r.t. $S(R)$ and some topology on $Ter^m(\Sigma)$. Then $S_E(R)$ is inconsistent w.r.t. that topology.*

Proof Take the sequence $t_n = t[x]_{f(n)}$. Then $t_n\ S(R)^{-1}\ t_{n+1}$ which is witnessed by context $t[\;]_{f(n)}$. As these contexts are eventually safe for any open neighbourhood of t we have $x = t_0\ S_E(R)\ t$. Therefore also $y\ S_E(R)\ t$ and so $x\ S_E(R)\ y$. □

Collapsing towers play a similar but not quite the same role as hypercollapsing terms. Originally, hypercollapsing terms were defined as those that have a hypercollapsing reduction sequence [20]; later this was modified [22] requiring all reducts of a term to have this property as well. The reason for the modification (which for orthogonal systems is immaterial) was to move towards a notion of *meaningless terms* [21,24] that also works for non-confluent systems. Collapsing towers indeed exhibit a hypercollapsing reduction sequence, so they are hypercollapsing w.r.t. the original definition, but there is no guarantee that this property is preserved by reduction, so they may not be hypercollapsing w.r.t. the new definition:

Example 21 Consider the iTRS with rules $F(x) \to x$ and $F(x) \to A$ w.r.t. metric d_∞. The term F^∞ is a collapsing tower, yet it rewrites to the normal form A and is therefore not hyper-collapsing by the new definition.

Generally, the mere presence of infinite terms can affect the equational theory, e.g. [3, Proposition 2.17] shows this for certain equational theories for the infinitary λ-calculus. The very presence of an infinite term in the term universe amounts to a uniformity property about the functions that make up that infinite term: trees made up from them converge regardless what appears low down the tree.

Considering other equivalences than $S_E(R)$, the inclusions

$$S_E(R) \subseteq 1_E(R) \subseteq T_E(R) \subseteq 2_E(R) \subseteq 3_E(R)$$

mean that the consistency of the later relations in the sequence implies the consistency of the earlier ones. Regarding the strictness of these inclusions we conjecture that this still applies to consistency, i.e. that there are (albeit increasingly contrived) examples in which the later equivalence is inconsistent and the earlier is not. For the first inclusion such an example is easy enough to construct:

Example 22 Consider the iTRS with rules $F(x) \to F(G(x))$ and $F(x) \to x$ w.r.t. term metric m with umms $F_m(\epsilon) = \epsilon$ and $G_m(\epsilon) = \epsilon/2$. Then the equivalence $S_E(\to_R)$ is consistent and $1_E(\to_R)$ is inconsistent. The inconsistency of $1_E(\to_R)$ follows from $F(x) \twoheadrightarrow_w x$ and $F(x) \twoheadrightarrow_w G^\infty$. $S_E(\to_R)$ is consistent, because (i) the collapsing tower F^∞ does not exist in $Ter^m(\Sigma)$, and (ii) the reduction from $F(x)$ to G^∞ is not strongly continuous.

The equivalence induced by metric models can become inconsistent for a term metric, even when the metric completion fails to produce any infinite terms:

Example 23 Consider the iTRS with the two collapsing rules $F(x) \to x$ and $G(x) \to x$. Set the term metric m such that $F_m(\epsilon) = \epsilon^2$ and $G_m(\epsilon) = \min(0.5, \epsilon)$. In this case $Ter(\Sigma) = Ter^m(\Sigma)$.

However, $d_m(F^n(G(x)), F^n(G(y))) = 2^{-n}$ and $F^n(G(x)) \to_R^* x$. Thus any metric model M can set the distance of $[\![x]\!]_M$ and $[\![y]\!]_M$ to be at most 2^{-n}, for any n. Hence $[\![x]\!]_M = [\![y]\!]_M$ in all metric models.

One can take several strategies to avoid inconsistency:

- Regard collapsing rules as “bad”.
- Use a variety of term metrics. After all, the cause of the inconsistency (for S_E) are collapsing towers — in other metrics those could be avoided, so that collapsing rules do not automatically lead to inconsistency. Not only that, an iTRS using another metric could be strongly converging even if it has collapsing rules. A case in point would be terminating and confluent finite TRSs, as these are special cases of iTRSs obtained from the trivial metric.

- Regard the majority of these equivalences as "too strong". This creates a semantic mismatch if we only look (as the literature has so far) at metric spaces as equational models, because the equivalence induced by these models is at least $3_E(\to_R)$.

If we take the last point we may consider even $S_E(\to_R)$ as too strong. A potentially interesting alternative is the equivalence $0_E(S(\to_R))$. The reason for this is two-fold.

- 0_E subsumes Σ^m_E, thus $0_E(S(\to_R))$ is a strong congruence relation;
- the construction joins topologically indistinguishable equivalence classes; in particular, any two R-collapsing towers are equivalent in $0_E(R)$.

Thus, viewing collapsing towers as "meaningless", this relation allows the arbitrary replacement of meaningless subterms across an infinite term, somewhat similar to $\sim_{hc}$. So it might be the case that reduction in orthogonal systems is confluent modulo $0_E(S(\to_R))$ as well.

7 Conclusion

Infinitary Rewriting deals with terms of potentially infinite size, and reductions of potentially infinite length. Working with these requires a machinery that can lift the finite to the infinite. In the literature of Infinitary Rewriting much of this lifting machinery is presented in an *ad hoc* way, defining various relations on infinite terms with their own purpose-built fully integrated machinery.

In this paper we have been looking at the generic side of this, uncovering the intrinsic closure principles that under-pin known constructions of transfinite rewrite relations and transfinite equivalences, such as weak and strong convergence. New is a generalisation of strong convergence to arbitrary topologies; this is also more elegant than the standard definition as it is based on a single principle. Also new is that strong convergence can indeed be decomposed into applying a closure operator to a single-step relation, which is not possible for the construction of weak convergence.

We also looked specifically at closure operators for equivalences, in particular equivalences that are closed under the same operators as rewriting is. We studied situations under which equivalences closed under strong sequences become inconsistent: for d_∞ in the presence of collapsing rules, more generally in the presence of "collapsing towers". Some of the closure operators on equivalences can be derived semantically from equational models. New in this paper is the notion of an equational topological model, for infinitary terms. This is largely inspired by the concept of an algebra of a monad in category theory, and takes a continuous interpretation map as a primitive rather than a derived concept. We show that any set of equations has an initial model, even if we restrict the model class to topological spaces that are T_0, T_1, T_2 or T_3. This can be extended to metric spaces, provided the interpretations are limited to non-expansive functions. Metric models are constructed in such a way that interpretations for infinite terms come for free.

References

1. Arnold, A., Nivat, M.: Metric interpretations of infinite trees and semantics of nondeterministic recursive programs. Theor. Comput. Sci. **11**(2), 181–205 (1980)
2. Bahr, P.: Partial order infinitary term rewriting and Böhm Trees. In: Lynch, C. (ed.) Proceedings of the 21st International Conference on Rewriting Techniques and Applications, Leibniz International Proceedings in Informatics (LIPIcs), vol. 6, pp. 67–84. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany (2010) doi:10.4230/LIPIcs.RTA.2010.67. http://drops.dagstuhl.de/opus/volltexte/2010/2645

3. Barendregt, H., Klop, J.W.: Applications of infinitary lambda calculus. Inf. Comput. **207**(5), 559–582 (2009). doi:10.1016/j.ic.2008.09.003
4. Barendregt, H.P.: The Lambda-Calculus, Its Syntax and Semantics. North-Holland, Amsterdam (1984)
5. Barr, M.: Terminal coalgebras in well-founded set theory. Theor. Comput. Sci. **114**, 299–315 (1993)
6. Davey, B.A., Priestley, H.A.: Introduction to Lattices and Order. Cambridge University Press, Cambridge (1990)
7. Dershowitz, N., Kaplan, S.: Rewrite, Rewrite, Rewrite, Rewrite, Rewrite, ... In: Principles of Programming Languages, pp. 250–259. ACM press (1989)
8. Dershowitz, N., Kaplan, S., Plaisted, D.: Rewrite, Rewrite, Rewrite, ... Theor. Comput. Sci. **83**(1), 71–96 (1991)
9. Endrullis, J., Grabmayer, C., Hendriks, D., Klop, J.W., Vrijer, R.: Proving infinitary normalization. In: Types for Proofs and Programs: International Conference, TYPES 2008 Torino, Italy, March 26–29, 2008 Revised Selected Papers, pp. 64–82. Springer, Berlin (2009)
10. Fox, R.: On topologies for function spaces. Bull. Am. Math. Soc. **51**, 429–432 (1945)
11. Gamelin, T.W., Greene, R.E.: Introduction to Topology. Dover Mineola, New York (1999)
12. Hahn, H.: Reelle Funktionen. Chelsea Publishing Company, New York (1948)
13. Hajnal, A., Hamburger, P.: Set Theory. London Mathematical Society, London (1999)
14. Isihara, A.: Algorithmic Term Rewriting Systems. Ph.D. thesis, Vrije Universiteit Amsterdam (2010)
15. Jänich, K.: Topology. Springer, Berlin (1984)
16. Kahrs, S.: Infinitary rewriting: Meta-theory and convergence. Acta Informatica **44**(2), 91–121 (2007)
17. Kahrs, S.: Modularity of convergence in infinitary rewriting. In: Treinen, R. (ed.) Rewriting Techniques and Applications, LNCS, vol. 5595, pp. 179–193. Springer, Berlin (2009)
18. Kahrs, S.: Infinitary rewriting: Foundations revisited. In: Lynch, C. (ed.) Proceedings of the 21st International Conference on Rewriting Techniques and Applications, Leibniz International Proceedings in Informatics (LIPIcs), vol. 6, pp. 161–176. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany (2010). doi:10.4230/LIPIcs.RTA.2010.161. http://drops.dagstuhl.de/opus/volltexte/2010/2651
19. Kahrs, S.: Modularity of convergence and strong convergence in infinitary rewriting. Log. Methods Comput Sci **6**(3), 27 (2010)
20. Kennaway, R., Klop, J.W., Sleep, R., de Vries, F.J.: Transfinite reductions in orthogonal term rewriting systems. Inf. Comput. **119**(1), 18–38 (1995)
21. Kennaway, R., van Oostrom, V., de Vries, F.J.: Meaningless terms in rewriting. J. Funct. Log. Program. **1**, 1–35 (1999)
22. Kennaway, R., de Vries, F.J.: Term Rewriting Systems, chap. Infinitary Rewriting. Cambridge University Press (2003)
23. Ketema, J.: Böhm-Like Rrees for Rewriting. Ph.D. thesis, Vrije Universiteit Amsterdam (2006)
24. Ketema, J.: Comparing Böhm-like trees. In: Treinen, R. (ed.) RTA, Lecture Notes in Computer Science, vol. 5595, pp. 239–254. Springer, Berlin (2009)
25. Kuratowski, K.: Topology. Academic Press, New York (1966)
26. Lüth, C.: Compositional term rewriting: An algebraic proof of Toyama's Theorem. In: Rewriting Techniques and Applications, pp. 261–275 (1996). LNCS 1103
27. MacLane, S.: Categories for the Working Mathematician. Springer, Berlin (1971)
28. Meinke, K., Tucker, J.: Universal algebra. In: Abramsky, S., Gabbay, D., Maibaum, T. (eds.) Handbook of Logic in Computer Science, vol. 1, pp. 189–411. Oxford University Press, Oxford (1992)
29. Potter, M.D.: Sets: An Introduction. Oxford University Press, Oxford (1990)
30. Sannella, D., Tarlecki, A.: Foundations of Algebraic Specification and Formal Software Development. Springer, Berlin (2012)
31. Simonsen, J.G.: Weak convergence and uniform normalization in infinitary rewriting. In: Lynch, C. (ed.) Proceedings of the 21st International Conference on Rewriting Techniques and Applications, Leibniz International Proceedings in Informatics (LIPIcs), vol. 6, pp. 311–324. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany (2010). doi:10.4230/LIPIcs.RTA.2010.311. http://drops.dagstuhl.de/opus/volltexte/2010/2660
32. Smyth, M.: Topology. In: Abramsky, S., Gabbay, D., Maibaum, T. (eds.) Handbook of Logic in Computer Science, vol. 1, pp. 641–762. Oxford University Press, Oxford (1992)
33. Steen, L.A., Seebach, J.A.: Counterexamples in Topology. Dover, New York (1995)
34. Thomas, W.: Automata on infinite objects. In: van Leeuwen, J. (ed.) Handbook of Theoretical Computer Science, vol. B, chap. 4, pp. 133–191. Elsevier, Amsterdam (1990)
35. Urysohn, P.: Zum metrizations problem. Math. Ann. **94**, 309–315 (1925)
36. Zantema, H.: Normalisation of infinite terms. In: RTA 2008, LNCS, vol. 5117, pp. 441–455. Springer, Berlin (2008)