1991

J.J.M.M. Rutten

Hereditarily-finite sets and complete metric spaces

Computer Science/Department of Software Technology Report CS-R9148 October

CWI, nationaal instituut voor onderzoek op het gebied van wiskunde en informatica

CWI is the research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a non-profit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands organization for scientific research (NWO).

Hereditarily Finite Sets and Complete Metric Spaces

J.J.M.M. Rutten
CWI
Kruislaan 413
1098 SJ Amsterdam
The Netherlands

Abstract

The collection of hereditarily finite sets is considered under the assumption of a nonstandard axiom system, in which the axiom of foundation has been replaced by a strong version of its negation, the anti-foundation axiom. An isomorphism is established between this collection of hereditarily finite sets and a complete metric space obtained as solution of a recursive domain equation (defined under the assumption of ZFC, i.e., without the anti-foundation axiom). Thus nonwellfounded hereditarily finite sets can be viewed as limits of Cauchy sequences of their wellfounded approximations.
1985 Mathematics Subject Classification: 68Q10, 68Q45.
1986 Computing Reviews Categories: D.3.1, F.3.2, F.3.3.
Key Words and Phrases: Nonwellfounded sets, anti-foundation axiom, bisimulation, complete metric spaces, hereditarily finite sets.
Note: Work partially supported by Basic Research Action Integration (3020).

1 Introduction

A **hereditarily finite** set is a set that first, is finite, second, of which all elements are finite, third, of which the elements of its elements are finite, and so on. Although this description is rather intuitive, we all are familiar with many examples of such sets: the natural numbers, defined as usual ($0 = \emptyset$ and $n + 1 = \{0, \ldots, n\}$), all are hereditarily finite.

Another way of trying to define this notion would be: A hereditarily finite set is a set that, first, is finite and second, the elements of which are hereditarily finite. This reflexive description suggests the following equation for the collection x of all hereditarily finite sets to satisfy:

$$x = \mathcal{P}_{<\omega}(x) \qquad (*)$$

Here $\mathcal{P}_{<\omega}(x)$ is the set of finite subsets of x. This equation can readily be seen to have a solution. Define $V_0 = \emptyset$, and given V_i, let $V_{i+1} = \mathcal{P}(V_i)$. Put HF $= \bigcup_i V_i$. Then HF satisfies $(*)$. In fact it is the smallest solution (under ordinary set inclusion). An interesting question now is the following: does it contain *all* hereditarily finite sets?

Report CS-R9148
ISSN 0169-118X
CWI
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

The answer to this question depends on the axiom system one is working with. If this is the usual Zermelo-Fraenkel set theory (ZFC) then the answer is affirmative: the Axiom of Foundation (FA) provides us with the principle of $\in$ induction, which can be straightforwardly used to show that HF is also the largest solution to $(*)$.

The existence of a set a satisfying $a = \{a\}$ is inhibited by the same FA: it is **nonwellfounded** in the sense that it contains an infinite descending $\in$ chain. Hence such a set cannot exist according to ZFC. Nevertheless, one can argue that it is just another example of a hereditarily finite set: the assumption that it is, is consistent with all other axioms of ZFC but for the FA. Clearly a is finite, and supposing a *is* hereditarily finite, also its elements are.

Recently, an alternative axiom system is receiving more and more attention, in which the FA is replaced by a strong version of its negation, a so-called **Anti-Foundation-Axiom** (AFA). This is mostly due to Peter Aczel ([Acz88]), who developed an intuitive theory and showed how nonwellfounded sets are of relevance to the study of computation and communication. Aczel's theory ZFA consists of ZFC - FA + AFA. An immediate consequence of the AFA is the existence of nonwellfounded sets. For instance, the above set a is in his approach an ordinary set like any other.

The formulation of the AFA uses the conception of sets as graphs. Every set x gives rise to a graph by taking as nodes the transitive closure of x and as (directed) edges all pairs (a, b) with $b \in a$. This graph is called the **canonical picture** of x. Note that wellfounded sets (which, by the way, still exist in ZFA) yield graphs of which all branches are finite (corresponding to the finiteness of descending $\in$ chains). Conversely, one might wonder whether there corresponds a set to every such (accessible pointed) graph. The AFA states that this is the case, and that moreover this set is unique. Next a notion that originally stems from the theory of computing is borrowed to describe an important property of $\mathcal{V}$, the collection of all sets under ZFA. It concerns the idea of **bisimulation,** which was originally introduced by Park ([Par81]). In $\mathcal{V}$ any two sets that are bisimilar (in a sense closely related to the original notion) are equal (and vice versa).

All this has a bearing on the equation $(*)$ above. In ZFA, it turns out to have *two* solutions, HF and in addition HF_1, the largest fixed point of $(*)$. It will contain HF as a subset and moreover nonwellfounded hereditarily finite sets like a.

The main characteristic of ZFA is that it allows sets to be nonwellfounded, hence countenancing self-membership. This makes it possible to give solutions to varying problems involving self-reference and circularity (like, e.g., described in [BE88]). In the theory of computing and more specifically, in the study of semantics of programming language semantics, this leads to transparent descriptions of recursive behaviour. (See the last chapter of [Acz88] and [Rut90].)

In the semantics of programming language semantics, infinite behaviour is traditionally described with the use of domain theory of some kind, in which the meaning of a recursive construct is obtained as the limit (in some sense) of the meanings of its finite approximations. Examples of such domains are Scott domains, complete partial orders and complete metric spaces.

A question already raised by Aczel himself is whether something similar could be done with nonwellfounded sets; that is, would it be possible to view these sets as limits of their wellfounded approximations?

A first solution to this problem is given by Mislove, Moss and Oles in their paper [MMO89]. The authors show how to interpret (hereditarily finite) nonwellfounded sets as fixed points of continuous transformations of an initial continuous algebra. This continuous algebra is obtained as the ideal completion of a preordered algebra closely related to HF. Its elements are called *protosets,* which can be seen as partially defined hereditarily finite sets. Intuitively, the work in [MMO89] amounts to the construction of a domain (which is later in their paper shown indeed to satisfy a domain equation) of finitely branching tree-like structures.

In this paper, we follow the same approach but now with the use of **complete metric spaces** as the basic mathematical tool, rather than the ordered structures mentioned above. The theory of solving reflexive domain equations in a category of complete metric spaces, as developed in [AR89] and initiated by [Niv79] and [dBZ82], is taken as a starting point. It gives for a large collection of domain equations the existence of a unique solution. This is applied to a variant of the above equation (*). The resulting domain contains (representations of commutative) unlabelled trees, called **processes,** and comes with a natural distance function (or metric): the distance between two trees is small if their finite approximations (truncations) are equal at great depth. An edge in such a process is related to an edge in the graph representations of hereditarily finite sets: it denotes the membership relation. Next a mapping is defined that associates with every hereditarily finite set a process, representing its canonical picture. Every nonwellfounded hereditarily finite set is mapped onto the limit of the Cauchy sequence consisting of the values of its wellfounded approximations. Thus this mapping is an ismorphism between HF_1 and a subset of the collection of processes.

The model of [MMO89] makes use of a domain that is built up directly from HF. The authors argue that this approach is preferable to starting with the solution of a domain equation straightaway, since it provides more intuition. Moreover they indicate that it is not completely clear to them what domain equation would actually be best.

Here we take the opposite approach by starting with the domain equation, which is a variant of (*) above, for which a complete metric space is taken as the solution. Interesting enough, the domain equation we use differs from both domain equations that are mentioned in [MMO89].

The metric domain is technically speaking simpler, and hence yields a more transparent model than its order-theoretic counterpart. This has three reasons. First, the use of metrics in a situation where a notion of *depth* exists, as is the case with the canonical pictures of hereditarily finite sets, leads to an intuitive model. Second, the theory of [AR89] yields immediately a solution for domain equations like (a variant of) (*) above. Third, the metric powerdomain constructor, which is used in that domain equation, is simpler than its order-theoretic counterparts.

However, the simplification thus obtained has its price: in the metric framework described above, one can only talk about completed elements; the notion of partial elements does not exist. This is opposed to the domain of [MMO89], where such partial elements do exist, the so-called protosets.

Finally we would like to mention some other related research. There is the work of Abramsky ([Abr88]), which is mentioned in [MMO89] but has not yet been published, and which compares several descriptions of finitary nonwellfounded sets. A similar remark applies to the work of Boffa ([Bof88]). (See [MMO89] for some discussion on the work of both authors.) Finally, a kind of overview paper on this subject is in preparation ([Mos91]).

Acknowledgement: As mentioned, this work closely follows the approach of [MMO89]

and owns therefore much to that paper.

2 Systems

In this section, the basic definitions concerning non-well-founded sets as developed in the first two chapters of [Acz88] are reviewed. (It closely follows a similar section in [MMO89].) For more motivation of the various definitions see [Acz88] or [BE88].

A **system** is a pair $S = \langle |S|, \xrightarrow[S]{} \rangle$, where $|S|$ is a set or class whose elements are called the **nodes** of S, and $\xrightarrow[S]{}$ is a binary **edge relation** on $|S|$. We frequently omit subscripts on arrows. If $x \rightarrow y$, then we call y a **child** of x. The family of children of each node of a system is required to form a set. A system S is **small** if $|S|$ is a set. An **accessible pointed graph**, or apg, is a triple $G = \langle |G|, \xrightarrow[G]{}, \mathrm{top}_G \rangle$ such that $\langle |G|, \xrightarrow[G]{} \rangle$ is a small system, top_G is a distinguished node of G, and every node of G is accessible from top_G. If G is an apg, denote the small system obtained by forgetting the distinguished node of G by $G^* = \langle |G|, \xrightarrow[G]{} \rangle$. If S is a system, and $x \in |S|$, then the apg $Sx = \langle |Sx|, \xrightarrow[Sx]{}, x \rangle$ is obtained by letting $|Sx|$ be the set of nodes of S accessible from x via $\xrightarrow[S]{}$, and by letting $\xrightarrow[Sx]{}$ be the restriction of $\xrightarrow[S]{}$ to $|Sx|$.

We mention three important examples of systems. First, the class V of all sets can be turned into a system $\mathcal{V}$ by taking as nodes the class V itself, and as edges the arrows $x \rightarrow y$ when $y \in x$. A second system, $\mathcal{S}$, is obtained by taking as nodes all accessible pointed graphs and by requiring $G \xrightarrow[\mathcal{S}]{} H$ iff $H = G^*n$ for some child n of top_G. Finally, let S be any system and x a node of S. Then we can form first the apg Sx and then forget the top to get the subsystem of S defined by x: $(Sx)^*$.

A **bisimulation** is a relation $\equiv$ on the nodes of a system such that $x \equiv y$ implies that for each child a of x there is some child b of y such that $a \equiv b$, and for each child b of y there is some child a of x such that $a \equiv b$. A bisimulation is not necessarily an equivalence relation. For every system S, there is a **maximal bisimulation** $\equiv_S$ on S which includes all bisimulation relations. The maximal bisimulation on a system is an equivalence relation. Two nodes x and y of a system S are **bisimilar** if $x \equiv_S y$.

Suppose S and T are systems. A map $f : |S| \rightarrow |T|$ is a **system map** from S to T provided that f preserves the sets of children. That is, for all nodes x of S,

$$\{f(y) \; : \; x \xrightarrow[S]{} y \} \; = \; \{z \; : \; f(x) \xrightarrow[T]{} z \}.$$

Every function $f : |S| \rightarrow |T|$ induces an equivalence relation $\equiv_f$ on its domain. Explicitly, $x \equiv_f y$ iff $f(x) = f(y)$. If $f : S \rightarrow T$ is a system map, then $\equiv_f$ is a bisimulation. A system map $f : S \rightarrow T$ is a **strongly extensional quotient** of S if f is surjective on nodes and $\equiv_f$ is exactly $\equiv_S$. Every system has a strongly extensional quotient. If $f : S \rightarrow T$ and $f' : S \rightarrow T'$ are two strongly extensional quotients of S, then T and T' are isomorphic systems.

A system S is called **strongly extensional** if two nodes of S are bisimilar if and only if they are equal; in other words, if $=_S$, equality on S, and $\equiv_S$, bisimilarity on S, are the same. This terminology can be explained by the observation that a strongly extensional system has a trivial strongly extensional quotient, namely itself.

If G is an apg, then a **decoration** of G is a system map $d : G^* \rightarrow \mathcal{V}$, and we say G is a **picture** of the set $d(\mathrm{top}_G)$. If x is a set, then the apg $\mathcal{V}x$ is called the **canonical**

picture of x because the inclusion function $d : |\mathcal{V}x| \to |\mathcal{V}|$ is a system map from $(\mathcal{V}x)^*$ to $\mathcal{V}$ such that $d(x) = x$.

3 The Anti-Foundation Axiom

The Anti-Foundation Axiom (**AFA**) states that for every apg G there is a unique system map $d_G : G^* \to \mathcal{V}$. An important consequence of the AFA is the fact that $\mathcal{V}$ is strongly extensional.

Theorem 3.1 *$\mathcal{V}$ is strongly extensional.*

Proof See [Acz88], Proposition 2.10. ⊣

Next consider the mapping $e : \mathcal{S} \to \mathcal{V}$ defined by, for all G, $e(G) = d_G(\mathrm{top}_G)$, where d_G is the unique decoration of G. The AFA implies that it is a strongly extensional quotient of $\mathcal{S}$.

Theorem 3.2 *The mapping $e : \mathcal{S} \to \mathcal{V}$ is*

(1) *a system map* (2) *a strongly extensional quotient*

Proof Part (1) of the theorem is trivial. For (2), one has to prove

$$\equiv_e = \equiv_{\mathcal{S}}$$

Trivially, we have $\equiv_{\mathcal{S}} \supseteq \equiv_e$. Conversely, let $G \equiv_{\mathcal{S}} H$. Because e is a system map this implies $e(G) \equiv_{\mathcal{V}} e(H)$. The fact that $\mathcal{V}$ is strongly extensional implies $e(G) = e(H)$. ⊣

For each function $f : X \to Y$ and each subclass Z of X, let $f[Z]$ denote the image of Z under f. The following definitions and theorem are introduced in [MMO89]; they will be used in the next section.

Definition If $S = \langle |S|, \underset{S}{\longrightarrow} \rangle$ is a system, then a **subsystem** of S is a system $T = \langle |T|, \underset{T}{\longrightarrow} \rangle$ where $|T| \subseteq |S|$ and $\underset{T}{\longrightarrow} = \underset{S}{\longrightarrow} \cap (|T| \times |T|)$. A **transitive subsystem** of S is a subsystem $T = \langle |T|, \underset{T}{\longrightarrow} \rangle$ also satisfying the property that, if $x \in |T|, y \in |S|$ and $x \underset{S}{\longrightarrow} y$, then $y \in |T|$.

Theorem 3.3 *If T is a transitive subsystem of S and $f : S \to S'$ is a system map, then the restriction $f|_T : T \to f[T]$ is a system map. If $f : S \to S'$ is a strongly extensional quotient, then so is $f|_T : T \to f[T]$.*

4 Hereditarily Finite Sets

A **hereditarily finite** set is a finite set, of which all the elements are finite, and so on. A way of defining this notion more precisely is to consider the following equation. The collection of all hereditarily finite sets should satisfy

$$x = \mathcal{P}_{<\omega}(x) \qquad (*)$$

Here $\mathcal{P}_{<\omega}(x)$ is the set of finite subsets of x. It is an instance of a so-called **class operator**, which takes a class x as an argument and yields another class $\mathcal{P}_{<\omega}(x)$.

Aczel showed for a large collection of class operators, namely the so-called set-continuous ones, that they have a smallest and a largest (with respect to set inclusion) fixed point. (See [Acz88]and [AM89].)

For our particular example of hereditarily finite sets, one need not resort to this general result. Assuming the axioms of set theory (notably Pairing, Union, Infinity, and some of the Replacement Axioms, but neither Foundation nor Choice), it is not difficult to show that the smallest fixed point of $(*)$, denoted by HF, exists. It can be constructed in the following way: Define $V_0 = \emptyset$, and given V_i, let $V_{i+1} = \mathcal{P}(V_i)$. Then $\mathrm{HF} = \bigcup_i V_i$.

In contrast to this, the largest fixed point of $(*)$, to be denoted by HF_1 cannot be shown to exist as a set without some Axiom of Foundation or Anti-Foundation. If we assume the usual Axiom of Foundation, we can straightforwardly show $\mathrm{HF}_1 = \mathrm{HF}$ by $\in$-induction. But without this axiom we have no principle of proof by $\in$-induction. In fact, under the AFA, HF is a proper subset of HF_1.

Definition An apg G is **finitely branching** if each node of G has finitely many children. $\mathcal{G}$ is the subsystem of $\mathcal{S}$ whose class of nodes consists of all finitely branching accessible pointed graphs.

Proposition 4.4 *Assume AFA. Then* HF_1 *is the set of all sets whose pictures are finitely branching accessible pointed graphs; i.e.,* $\mathrm{HF}_1 = e[\,|\mathcal{G}|\,]$.

Proof This result is implicit in [Acz88]. A direct proof can be found in [MMO89]. It shows that $e[\,|\mathcal{G}|\,]$ is a set rather than a proper class, that $e[\,|\mathcal{G}|\,]$ is a fixed point of $(*)$, and that every other fixed point of $(*)$ is contained in $e[\,|\mathcal{G}|\,]$. ⊣

Theorem 4.5 HF_1 *is a strongly extensional quotient of* $\mathcal{G}$.

Proof By Theorem 3.3 the restriction of $e : \mathcal{S} \to \mathcal{V}$ to $\mathcal{G}$, which is a strongly extensional quotient of $\mathcal{S}$, is again a strongly extensional quotient. Now the theorem is implied by Proposition 4.4. ⊣

5 Metric Processes

Next we introduce a complete metric space, P, that will serve as a co-domain for a system map for $\mathcal{G}$, to be constructed in the next section. The domain P is defined as the solution of a reflexive domain equation. (Everything one needs to know about metric spaces can be found in the first appendix. Those interested in the details of solving domain equations with the help of complete metric spaces are referred to the second appendix, which gives a brief abstract of [AR89].)

Definition Let (P, d) be the unique complete metric space satisfying

$$P \cong \mathcal{P}_{co}(id_{1/2}(P))$$

Here $\mathcal{P}_{co}(id_{1/2}(P))$ is the collection of all compact subsets of $id_{1/2}(P)$; the latter space is like P but with a different metric, namely, $1/2 \cdot d$. The elements of P are called **processes,** and will be indicated by the symbols x, y, X, Y, p, q. The symbol $\cong$ should be read as "is isometric (isomorphic) to". That is, there exist unique distance-preserving mappings

$$\mu : P \longrightarrow \mathcal{P}_{co}(id_{1/2}(P))$$

$$\nu : \mathcal{P}_{co}(id_{1/2}(P)) \longrightarrow P$$

such that

$$\mu \circ \nu = id, \quad \nu \circ \mu = id$$

Usually we omit μ and ν and identify P and $\mathcal{P}_{co}(id_{1/2}(P))$. The metric $d : P \times P \longrightarrow [0,1]$ is defined implicitly by the above equation; it satisfies, for all $X, Y \in P$,

$$d(X, Y) = (1/2 \cdot d)_H(\mu(X), \mu(Y))$$

Here $(1/2 \cdot d)_H$ is the so-called Hausdorff metric on subsets of P, induced by the metric $1/2 \cdot d$ on P.

The metric d on P has a very intuitive characterization.

Lemma 5.6 *Let for $X \in P$ and $n \geq 0$ the n-th truncation $X[n]$ of X be defined by*

$$X[0] = \emptyset$$

$$X[n+1] = \{x[n] : x \in X\}$$

Then for all $X, Y \in P$,

$$d(X, Y) = 2^{-sup\{k : X[k] = Y[k]\}}$$

Proof For a formal proof one has to be precise about the use of μ and ν. The definition of $X[n+1]$ should actually be

$$X[n+1] = \nu(\{x[n] : x \in \mu(X)\})$$

Now the lemma follows from the defining equation for d. ⊣

Example

As usual let the natural numbers be defined by

$$0 = \emptyset, \quad n+1 = \{0, \ldots, n\}$$

Then for all n, k,

$$n[k] = \begin{cases} k & \text{if } k \leq n \\ n & \text{if } k > n \end{cases}$$

as can readily be shown by induction. Hence for all n, m,

$$d(n, m) = \begin{cases} 2^{-min\{n,m\}} & \text{if } n \neq m \\ 0 & \text{if } n = m \end{cases}$$

The metric space P is a system by taking as edge relation

$$p \xrightarrow[P]{} q =^{def} q \in \mu(p)$$

(We shall, again, often omit μ and write $q \in p$.) The system $(P, \xrightarrow[P]{})$, again indicated by P, has the interesting property that it is strongly extensional.

Theorem 5.7 *P is strongly extensional, i.e., for all $p, q \in P$,*

$$p \equiv_P q \textit{ if and only if } p = q$$

Proof Let $p, q \in P$ with $p \equiv_P q$. We show that $p = q$, the reverse implication being trivial. Let $R \subseteq P \times P$ be a bisimulation relation with pRq. Define

$$\epsilon = \sup\{d(X, Y) : X, Y \in P \wedge XRY\}$$

We prove that $\epsilon \leq 1/2 \cdot \epsilon$, which implies $\epsilon = 0$. Hence $d(p, q) = 0$ and thus $p = q$.

Let $x, y \in P$ with XRY. By the definition of the metric d on P, which implies that d equals $(d_{id_{1/2}(P)})_H$, it is sufficient to show that, for all $x \in X$, $d(x, Y) \leq 1/2 \cdot \epsilon$. Let $x \in X$. Because XRY there exists $y \in Y$ such that xRy. We have

$$d(x, Y) = \inf_{y' \in Y}\{d_{id_{1/2}(P)}(x, y')\} \leq 1/2 \cdot d(x, y) \leq 1/2 \cdot \epsilon$$

Hence $\epsilon \leq 1/2 \cdot \epsilon$. $\dashv$

6 Processes Associated to Graphs

Recall that $\mathcal{G}$ is the subsystem of $\mathcal{S}$ consisting of all finitely branching apg's (accessible pointed graphs). In this section, we shall associate to each finitely branching apg G (i.e., $G \in |\mathcal{G}|$) a process in P. To this end, first a sequence $(n_i)_i$ of processes is associated to every node n in G.

Definition Let $G \in |\mathcal{G}|$ and let n be a node in G. We define for all nodes at the same time a sequence $(n_i)_i$ of processes as follows:

$$n_0 = \emptyset, \quad n_{i+1} = \{m_i : n \longrightarrow m\}$$

(Here $n \longrightarrow m$ means there is an edge from n to m.)

Lemma 6.8 *The sequence $(n_i)_i$ is Cauchy.*

Proof The following fact can be easily proved: for all i and k,

$$n_i[k] = \begin{cases} n_k & \text{if } k \leq i \\ n_i & \text{if } k > i \end{cases}$$

Together with Lemma 5.6 this implies, for all i and j, $d(n_i, n_j) \leq 2^{-min\{i,j\}}$. ⊣

This lemma justifies the following definition, specifying a function that associates to every finitely branching apg a process as follows.

Definition Let $\delta : |\mathcal{G}| \to \mathcal{P}$ be defined, for $G \in |\mathcal{G}|$, by

$$\delta(G) = \lim_{i \to \infty} (top_G)_i$$

The mapping δ has the following properties.

Theorem 6.9 *The mapping $\delta : |\mathcal{G}| \to \mathcal{P}$ is*

(1) *a system map* (2) *a strongly extensional quotient*

Proof We identify apg's $G \in |\mathcal{G}|$ with their top nodes n.

(1) **δ is a system map:** We have to show

$$\{\delta(m) : n \underset{\mathcal{G}}{\longrightarrow} m\} = \{p : \delta(n) \underset{P}{\longrightarrow} p\}$$

This equality follows from the two inclusions $\subseteq$ and $\supseteq$, which we show next.

The following general fact can be conveniently used. Let (M, d) be a complete metric space. Consider $(\mathcal{P}_{nc}(M), d_H)$, where d_H is the Hausdorff metric induced by d. Let $(X_i)_i$ be a Cauchy sequence in $\mathcal{P}_{nc}(M)$. Then

$$\lim_{i \to \infty} X_i = \{\lim_{i \to \infty} x_i : \ x_i \in X_i \text{ and } (x_i)_i \text{ is a Cauchy sequence in } M\}$$

In particular, we have

$$\delta(n)(= \lim_{i \to \infty} n_i) =^{(*)} \{\lim_{i \to \infty} x_i : \ x_i \in n_i \text{ and } (x_i)_i \text{ is a Cauchy sequence in } P\}$$

$\subseteq$: We have

$$\begin{aligned} \{\delta(m) : n \underset{\mathcal{G}}{\longrightarrow} m\} &= \{\lim_{i \to \infty} m_i : \ n \underset{\mathcal{G}}{\longrightarrow} m\} \\ &= \{\lim_{i \to \infty} m_{i-1} : \ m_{i-1} \in n_i\} \\ &\subseteq \{\lim_{i \to \infty} x_i : \ x_i \in n_i \text{ and } (x_i)_i \text{ is a Cauchy sequence in } P\} \\ &= \text{(by (*)) } \delta(n) \\ &= \{p : \delta(n) \underset{P}{\longrightarrow} p\} \end{aligned}$$

$\supseteq$: We show: if $p \in \delta(n)$ then there is a node m such that $n \xrightarrow[\mathcal{G}]{} m$ and $\delta(m) = p$. Let $p \in \delta(n)$. By (*) there exists a Cauchy sequence $\lim_{i\to\infty} x_i$ with $x_i \in n_i$ and $p = \lim_{i\to\infty} x_i$. Since we consider finitely branching graphs, n has only finitely many sons $n \xrightarrow[\mathcal{G}]{} m$, say $\{m_1, \ldots, m_k\}$. Because $n_i = \{(m_1)_{i-1}, \ldots, (m_k)_{i-1}\}$, and $x_i \in n_i$ for all i, there exists l, $1 \leq l \leq k$, such that m_l occurs infinitely often in $\lim_{i\to\infty} x_i$. Let $f : \mathbf{N} \to \mathbf{N}$ be a monotonic function such that $x_{f(i)} = (m_l)_{f(i)-1}$, for all i. Then

$$p = \lim_{i\to\infty} x_i = \lim_{i\to\infty} x_{f(i)} = \lim_{i\to\infty} (m_l)_{f(i)-1} = \delta(m_l)$$

and $\delta(n) \xrightarrow[P]{} \delta(m_l)$.

(2): δ **is a strongly extensional quotient:** We have to show $\equiv_{\mathcal{G}} = \equiv_{\delta}$. Trivially, we have $\equiv_{\mathcal{G}} \supseteq \equiv_{\delta}$. Conversely, let $n \equiv_{\mathcal{G}} m$. Because δ is a system map this implies $\delta(n) \equiv_P \delta(m)$. Now the fact that P is strongly extensional (Theorem 5.7) implies $\delta(n) = \delta(m)$. ⊣

Finally, we can formulate the main result of our paper.

Theorem 6.10 HF_1 *and* $\delta[\,|\mathcal{G}|\,]$ *are isomorphic.*

Proof This is immediate from the fact that both HF_1 and $\delta[\,|\mathcal{G}|\,]$ are a strongly extensional quotient of $\mathcal{G}$ (Theorem 4.5 and Theorem 6.9). ⊣

References

[Abr88] S. Abramsky. A cook's tour of the finitary non-well-founded sets. Technical report, 1988. Unpublished notes.

[Acz88] P. Aczel. *Non-well-founded sets.* Number 14 in CSLI Lecture Notes. 1988.

[AM89] P. Aczel and N. Mendler. A final coalgebra theorem. In D.H. Pitt, D.E. Rye-heard, P. Dybjer, A. M. Pitts, and A. Poigne, editors, *Proceedings category theory and computer science*, Lecture Notes in Computer Science, pages 357–365, 1989.

[AR89] P. America and J.J.M.M. Rutten. Solving reflexive domain equations in a category of complete metric spaces. *Journal of Computer and System Sciences*, 39(3):343–375, 1989.

[BE88] J. Barwise and J. Etchemendy. *The Liar: An Essay in Truth and Circularity.* Oxford University Press, 1988.

[Bof88] M. Boffa. Finite approximations of sets. Technical report, 1988. Unpublished notes.

[dBZ82] J.W. de Bakker and J.I. Zucker. Processes and the denotational semantics of concurrency. *Information and Control*, 54:70–120, 1982.

[Eng77] R. Engelking. *General Topology.* Polish Scientific Publishers, 1977.

[Mac71] S. MacLane. *Categories for the Working Mathematician*, volume 5 of *Graduate Texts in Mathematics.* Springer-Verlag, 1971.

[MMO89] M.W. Mislove, L.S. Moss, and F.J. Oles. Non-well-founded sets obtained from ideal fixed points. In *Proc. of the Fourth IEEE Symposium on Logic in Computer Science*, pages 263–272, 1989. To appear in Information and Computation.

[Mos91] L.S. Moss. Personal communication. Technical report, 1991.

[Niv79] M. Nivat. Infinite words, infinite trees, infinite computations. In J.W. de Bakker and J. van Leeuwen, editors, *Foundations of Computer Science III, Part 2*, volume 109 of *Math. Centre Tracts*, pages 3–52, 1979.

[Par81] D.M.R. Park. Concurrency and automata on infinite sequences. In *Proceedings 5th GI conference*, volume 104 of *Lecture Notes in Computer Science*, pages 15–32. Springer-Verlag, 1981.

[Rut90] J.J.M.M. Rutten. Nonwellfounded sets and programming language semantics. Technical Report CS-R9063, Centre for Mathematics and Computer Science, Amsterdam, 1990. To appear in the proceedings of Mathematical Foundations of Programming Semantics, Pittsburgh, March 1991.

A Metric Spaces

We assume the following notions to be known (the reader might consult [Eng77]): metric space, ultra-metric space, complete (ultra-)metric space, continuous function, closed set, compact set. In our definition the distance between two elements of a metric space is always between 0 and 1, inclusive.

An arbitrary set A can be supplied with a metric d_A, called the *discrete* metric, defined by

$$d_A(x,y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y \end{cases}$$

Now (A, d_A) is a metric space (it is even an ultra-metric space).

Let (M_1, d_1) and (M_2, d_2) be two complete metric spaces. A function $f : M_1 \rightarrow M_2$ is called *non-expansive* if for all $x, y \in M_1$

$$d_2(f(x), f(y)) \leq d_1(x,y)$$

The set of all non-expansive functions from M_1 to M_2 is denoted by $M_1 \xrightarrow{1} M_2$. A function $f : M_1 \rightarrow M_2$ is called *contracting* (or a *contraction*) if there exists an $\epsilon < 1$ such that for all $x, y \in M_1$

$$d_2(f(x), f(y)) \leq \epsilon \cdot d_1(x,y)$$

(Non-expansive functions and contractions are always continuous.)

The following fact is known as Banach's theorem: Let (M, d) be a complete metric space and $f : M \rightarrow M$ a contraction. Then f has a unique fixed point, that is, there exists a unique $x \in M$ such that $f(x) = x$. This x can be obtained by taking the limit of $f^n(x_0)$ for any arbitrary $x_0 \in M$ (where $f^0(y) = y$ and $f^{n+1}(y) = f(f^n(y))$).

We call M_1 and M_2 *isometric* (notation: $M_1 \cong M_2$) if there exists a bijective mapping $f : M_1 \rightarrow M_2$ such that for all $x, y \in M_1$

$$d_2(f(x), f(y)) = d_1(x,y)$$

Definition Let (M, d), $(M_1, d_1), \ldots, (M_n, d_n)$ be metric spaces.

1. We define a metric d_F on the set $M_1 \rightarrow M_2$ of all functions from M_1 to M_2 as follows: For every $f_1, f_2 \in M_1 \rightarrow M_2$ we put

   $$d_F(f_1, f_2) = \sup_{x \in M_1} \{d_2(f_1(x), f_2(x))\}$$

   This supremum always exists since the values taken by our metrics are always between 0 and 1. The set $M_1 \xrightarrow{1} M_2$ is a subset of $M_1 \rightarrow M_2$, and a metric on $M_1 \xrightarrow{1} M_2$ can be obtained by taking the restriction of the corresponding d_F.

2. With $M_1 \bar{\cup} \cdots \bar{\cup} M_n$ we denote the *disjoint union* of $M_1, \ldots, M_n$, which can be defined as $\{1\} \times M_1 \cup \cdots \cup \{n\} \times M_n$. We define a metric d_U on $M_1 \bar{\cup} \cdots \bar{\cup} M_n$ as follows: For every $x, y \in M_1 \bar{\cup} \cdots \bar{\cup} M_n$,

   $$d_U(x,y) = \begin{cases} d_j(x,y) & \text{if } x, y \in \{j\} \times M_j, 1 \leq j \leq n \\ 1 & \text{otherwise} \end{cases}$$

   If no confusion is possible we often write $\cup$ rather than $\bar{\cup}$.

3. We define a metric d_P on the Cartesian product $M_1 \times \cdots \times M_n$ by the following clause: For every $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in M_1 \times \cdots \times M_n$,

$$d_P((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \max_i \{ d_i(x_i, y_i) \}$$

4. Let $\mathcal{P}_{cl}(M) = \{ X : X \subseteq M \wedge X \text{ is closed} \}$. We define a metric d_H on $\mathcal{P}_{cl}(M)$, called the *Hausdorff distance*, as follows: For every $X, Y \in \mathcal{P}_{cl}(M)$,

$$d_H(X, Y) = \max\{ \sup_{x \in X} \{ d(x, Y) \}, \sup_{y \in Y} \{ d(y, X) \} \}$$

where $d(x, Z) = \inf_{z \in Z} \{ d(x, z) \}$ for every $Z \subseteq M$, $x \in M$. (We use the convention that $\sup \emptyset = 0$ and $\inf \emptyset = 1$.) The spaces $\mathcal{P}_{co}(M) = \{ X : X \subseteq M \wedge X \text{ is compact} \}$ and $\mathcal{P}_{nc}(M) = \{ X : X \subseteq M \wedge X \text{ is non-empty and compact} \}$ are supplied with a metric by taking the restriction of d_H.

5. For any real number ϵ with $0 < \epsilon \leq 1$ we define

$$\mathrm{id}_\epsilon((M, d)) = (M, d')$$

where $d'(x, y) = \epsilon \cdot d(x, y)$, for every x and y in M.

Proposition A.11 *Let (M, d), $(M_1, d_1), \ldots, (M_n, d_n)$, d_F, d_U, d_P and d_H be as in Definition A and suppose that (M, d), $(M_1, d_1), \ldots, (M_n, d_n)$ are complete. Then*

$$(M_1 \rightarrow M_2, d_F) \quad (M_1 \xrightarrow{1} M_2, d_F) \tag{a}$$

$$(M_1 \bar{\cup} \cdots \bar{\cup} M_n, d_U) \tag{b}$$

$$(M_1 \times \cdots \times M_n, d_P) \tag{c}$$

$$(\mathcal{P}_{cl}(M), d_H) \quad (\mathcal{P}_{co}(M), d_H) \quad (\mathcal{P}_{nc}(M), d_H) \tag{d}$$

$$\mathrm{id}_\epsilon((M, d)) \tag{e}$$

are complete metric spaces. If (M, d) and (M_i, d_i) are all ultra-metric spaces, then so are these composed spaces. (Strictly speaking, for the completeness of $M_1 \rightarrow M_2$ and $M_1 \xrightarrow{1} M_2$ we do not need the completeness of M_1. The same holds for the ultra-metric property.)

Whenever in the sequel we write $M_1 \rightarrow M_2$, $M_1 \xrightarrow{1} M_2$, $M_1 \bar{\cup} \cdots \bar{\cup} M_n$, $M_1 \times \cdots \times M_n$, $\mathcal{P}_{cl}(M)$, $\mathcal{P}_{co}(M)$, $\mathcal{P}_{nc}(M)$, or $\mathrm{id}_\epsilon(M)$, we mean the metric space with the metric defined above.

The proofs of Proposition A.11(a), (b), (c), and (e) are straightforward. Part (d) is more complex. It can be proved with the help of the following characterization of the completeness of $(\mathcal{P}_{cl}(M), d_H)$.

Proposition A.12 *Let $(\mathcal{P}_{cl}(M), d_H)$ be as in Definition A. Let $(X_i)_i$ be a Cauchy sequence in $\mathcal{P}_{cl}(M)$. We have*

$$\lim_{i \rightarrow \infty} X_i = \{ \lim_{i \rightarrow \infty} x_i : x_i \in X_i, (x_i)_i \text{ a Cauchy sequence in } M \}$$

Proofs of Propositions A.11(d) and A.12 can be found in, for instance, [Eng77]. The proofs are also repeated in [dBZ82].

A Metric spaces as solutions to domain equations

We show how one can use metric spaces to solve so-called ***reflexive domain equations*** of, e.g., the following form:

$$P \cong F(P)$$

(The symbol $\cong$ is defined below; it says that there is a bijection from P to $F(P)$ that respects the metric defined on the spaces.) Here $F(P)$ is an expression built from P and a number of standard constructions on metric spaces (also to be formally introduced shortly). A few examples are

$$P \cong A \cup (B \times P) \quad (1)$$

$$P \cong A \cup \mathcal{P}_{co}(B \times P) \quad (2)$$

$$P \cong A \cup (B \to P) \quad (3)$$

where A and B are given fixed complete metric spaces. De Bakker and Zucker have first described how to solve these equations in a metric setting [dBZ82]. Roughly, their approach amounts to the following: In order to solve $P \cong F(P)$ they define a sequence of complete metric spaces $(P_n)_n$ by: $P_0 = A$ and $P_{n+1} = F(P_n)$, for $n > 0$, such that $P_0 \subseteq P_1 \subseteq \cdots$. Then they take the *metric completion* of the union of these spaces P_n, say $\bar{P}$, and show: $\bar{P} \cong F(\bar{P})$. In this way they are able to solve equations (1), (2) and (3) above.

There is one type of equation for which this approach does not work, namely,

$$P \cong A \cup (P \xrightarrow{1} G(P)) \quad (4)$$

in which P occurs at the *left* side of a function space arrow, and $G(P)$ is an expression possibly containing P. This is due to the fact that it is not always the case that $P_n \subseteq F(P_n)$.

In [AR89] the above approach is generalized in order to overcome this problem. Moreover, it provides a more precise description of the solutions of domain equations and their metrics.

The family of complete metric spaces is made into a *category* $\mathcal{C}$ by providing some additional structure. (For an extensive introduction to category theory we refer the reader to [Mac71].) Then the expression F is interpreted as a *functor* $F : \mathcal{C} \to \mathcal{C}$ which is (in a sense) *contracting*. It is proved that a generalized version of Banach's theorem (see below) holds, i.e., that contracting functors have a fixed point (up to isometry). Such a fixed point, satisfying $P \cong F(P)$, is a solution of the domain equation.

We shall now give a quick overview of these results, omitting many details and all proofs. For a full treatment we refer the reader to [AR89]. The basic definitions and facts of metric topology that we shall need can be found in the first appendix.

We introduce a category of complete metric spaces and some basic definitions, after which a categorical fixed point theorem will be formulated.

Definition Let $\mathcal{C}$ denote the category that has complete metric spaces for its objects. The arrows ι in $\mathcal{C}$ are defined as follows: Let M_1, M_2 be complete metric spaces. Then $M_1 \to^{\iota} M_2$ denotes a pair of maps $M_1 \rightleftarrows^{i}_{j} M_2$, satisfying the following properties:

1. i is an isometric embedding,

2. j is non-expansive,

3. $j \circ i = \mathrm{id}_{M_1}$.

(We sometimes write $\langle i \rangle j$ for ι.) Composition of the arrows is defined in the obvious way.

We can consider M_1 as an approximation to M_2: In a sense, the set M_2 contains more information than M_1, because M_1 can be isometrically embedded into M_2. Elements in M_2 are approximated by elements in M_1. For an element $m_2 \in M_2$ its (best) approximation in M_1 is given by $j(m_2)$. Clause 3 states that M_2 is a consistent extension of M_1.

Definition For every arrow $M_1 \rightarrow^{\iota} M_2$ in $\mathcal{C}$ with $\iota = \langle i \rangle j$ we define

$$\delta(\iota) = d_{M_2 \rightarrow M_1}(i \circ j, \mathrm{id}_{M_2}) \quad (= \sup_{m_2 \in M_2} \{d_{M_2}(i \circ j(m_2), m_2)\})$$

This number can be regarded as a measure of the quality with which M_2 is approximated by M_1: the smaller $\delta(\iota)$, the better M_2 is approximated by M_1.

Increasing sequences of metric spaces are generalized as follows:

Definition We call a sequence $(D_n, \iota_n)_n$ of complete metric spaces and arrows a *tower* whenever we have that $\forall n \in \mathbf{N}\, D_n \rightarrow^{\iota_n} D_{n+1} \in \mathcal{C}$. The sequence $(D_n, \iota_n)_n$ is called a *converging tower* when the following condition is also satisfied:

$$\forall \epsilon > 0 \, \exists N \in \mathbf{N} \, \forall m > n \geq N \; \delta(\iota_{nm}) < \epsilon$$

where $\iota_{nm} = \iota_{m-1} \circ \cdots \circ \iota_n : D_n \rightarrow D_m$.

A special case of a converging tower is a tower $(D_n, \iota_n)_n$ satisfying, for some ϵ with $0 \leq \epsilon < 1$,

$$\forall n \in \mathbf{N} \; \delta(\iota_{n+1}) \leq \epsilon \cdot \delta(\iota_n)$$

Note that

$$\begin{aligned} \delta(\iota_{nm}) &\leq \delta(\iota_n) + \cdots + \delta(\iota_{m-1}) \\ &\leq \epsilon^n \cdot \delta(\iota_0) + \cdots + \epsilon^{m-1} \cdot \delta(\iota_0) \\ &\leq \frac{\epsilon^n}{1-\epsilon} \cdot \delta(\iota_0) \end{aligned}$$

We shall now generalize the technique of forming the metric *completion* of the union of an increasing sequence of metric spaces by proving that, in $\mathcal{C}$, every converging tower has an *initial cone*. The construction of such an initial cone for a given tower is called the *direct limit* construction. Before we treat this direct limit construction, we first give the definition of a cone and an initial cone.

Definition Let $(D_n, \iota_n)_n$ be a tower. Let D be a complete metric space and $(\gamma_n)_n$

a sequence of arrows. We call $(D, (\gamma_n)_n)$ a *cone* for $(D_n, \iota_n)_n$ whenever the following condition holds:

$$\forall n \in \mathbf{N}\, D_n \rightarrow^{\gamma_n} D \in \mathcal{C} \wedge \gamma_n = \gamma_{n+1} \circ \iota_n$$

Definition A cone $(D, (\gamma_n)_n)$ for a tower $(D_n, \iota_n)_n$ is called *initial* whenever for every other cone $(D', (\gamma'_n)_n)$ for $(D_n, \iota_n)_n$ there exists a unique arrow $\iota : D \rightarrow D'$ in $\mathcal{C}$ such that:

$$\forall n \in \mathbf{N}\, \iota \circ \gamma_n = \gamma'_n$$

Definition Let $(D_n, \iota_n)_n$, with $\iota_n = \langle i_n \rangle j_n$, be a converging tower. The *direct limit* of $(D_n, \iota_n)_n$ is a cone $(D, (\gamma_n)_n)$, with $\gamma_n = \langle g_n \rangle h_n$, that is defined as follows:

$$D = \{ (x_n)_n : \forall n \geq 0\, x_n \in D_n \wedge j_n(x_{n+1}) = x_n \}$$

is equipped with a metric d_D defined by

$$d_D((x_n)_n, (y_n)_n) = \sup\{ d_{D_n}(x_n, y_n) \}$$

for all $(x_n)_n$ and $(y_n)_n \in D$. The mapping $g_n : D_n \rightarrow D$ is defined by $g_n(x) = (x_k)_k$, where

$$x_k = \begin{cases} j_{kn}(x) & \text{if } k < n \\ x & \text{if } k = n \\ i_{nk}(x) & \text{if } k > n \end{cases}$$

and $h_n : D \rightarrow D_n$ is defined by $h_n((x_k)_k) = x_n$.

Lemma A.13 *The direct limit of a converging tower (as defined in Definition A) is an initial cone for that tower.*

As a category-theoretic equivalent of a contracting function on a metric space, we have the following notion of a *contracting functor* on $\mathcal{C}$.

Definition We call a functor $F : \mathcal{C} \rightarrow \mathcal{C}$ contracting whenever the following holds: There exists an ϵ, with $0 \leq \epsilon < 1$, such that, for all $D \rightarrow^{\iota} E \in \mathcal{C}$,

$$\delta(F(\iota)) \leq \epsilon \cdot \delta(\iota)$$

A contracting function on a complete metric space is continuous, so it preserves Cauchy sequences and their limits. Similarly, a contracting functor preserves converging towers and their initial cones:

Lemma A.14 *Let $F : \mathcal{C} \rightarrow \mathcal{C}$ be a contracting functor, let $(D_n, \iota_n)_n$ be a converging tower with an initial cone $(D, (\gamma_n)_n)$. Then $(F(D_n), F(\iota_n))_n$ is again a converging tower with $(F(D), (F(\gamma_n))_n)$ as an initial cone.*

Theorem A.15 *Let F be a contracting functor $F : \mathcal{C} \to \mathcal{C}$ and let $D_0 \to^{\iota_0} F(D_0) \in \mathcal{C}$. Let the tower $(D_n, \iota_n)_n$ be defined by $D_{n+1} = F(D_n)$ and $\iota_{n+1} = F(\iota_n)$ for all $n \geq 0$. This tower is converging, so it has a direct limit $(D, (\gamma_n)_n)$. We have $D \cong F(D)$.*

In [AR89] it is shown that contracting functors that are moreover contracting on all *hom-sets* (the sets of arrows in $\mathcal{C}$ between any two given complete metric spaces) have *unique* fixed points (up to isometry). It is also possible to impose certain restrictions upon the category $\mathcal{C}$ such that every contracting functor on $\mathcal{C}$ has a unique fixed point.

Let us now indicate how this theorem can be used to solve Equations (1) to (4) above. We define

$$F_1(P) = A \cup \mathrm{id}_{1/2}(B \times P) \tag{5}$$

$$F_2(P) = A \cup \mathcal{P}_{co}(B \times \mathrm{id}_{1/2}(P)) \tag{6}$$

$$F_3(P) = A \cup (B \to \mathrm{id}_{1/2}(P)) \tag{7}$$

If the expression $G(P)$ in Equation (4) is, for example, equal to P, then we define F_4 by

$$F_4(P) = A \cup \mathrm{id}_{1/2}(P \xrightarrow{1} P) \tag{8}$$

Note that the definitions of these functors specify, for each metric space (P, d_P), the metric on $F(P)$ *implicitly* (see Definition A).

Now it is easily verified that F_1, F_2, F_3, and F_4 are contracting functors on $\mathcal{C}$. Intuitively, this is a consequence of the fact that in the definitions above each occurrence of P is preceded by a factor $\mathrm{id}_{1/2}$. Thus these functors have a fixed point, according to Theorem A.15, which is a solution for the corresponding equation. (We often omit the factor $\mathrm{id}_{1/2}$ in the reflexive domain equations, assuming that the reader will be able to fill in the details.)

In [AR89] it is shown that functors like F_1 to F_4 are also contracting on hom-sets, which guarantees that they have *unique* fixed points (up to isometry).

The results above hold for complete *ultra-metric* spaces too, which can be easily verified.