Constructively Finite?

Arnaud Spiwack, Thierry Coquand

CONTRIBUCIONES CIENTÍFICAS
EN HONOR DE MIRIAN ANDRÉS GÓMEZ
(Laureano Lambán, Ana Romero y Julio Rubio, editores),
Servicio de Publicaciones, Universidad de La Rioja,
Logroño, Spain, 2010.

CONSTRUCTIVELY FINITE?

THIERRY COQUAND AND ARNAUD SPIWACK

Este artículo está dedicado a Mirian Andrés Gómez, con quien pasé momentos inolvidables y a quien extrañaré siempre.
– Arnaud Spiwack

Resumen. Este artículo explora varias propiedades de conjuntos "à la Bishop", que serían equivalentes en ZFC a la propiedad de ser finito. Una de ellas es nueva. Aprovechamos esta oportunidad para considerar detenidamente las diferentes propiedades de los conjuntos finitos, y cómo se relacionan (constructivamente) unas con otras. Cerramos el artículo presentando algunos algoritmos bien conocidos sobre estructuras finitas, y describiendo qué tipo de finitud requieren.

Abstract. This articles plays with several properties of Bishop sets which would be equivalent in ZFC to being finite. One of which is new. We take this oportunity to closely consider the different properties of finite sets and how they (constructively) relate to one another. We close this article by presenting a few well-known algorithms on finite structures, and describe which kind of finiteness they require.

1. Introduction

1.1. Foreword.

We shall try, in this article, to shake two rather common beliefs. First that constructive mathematics offer little extra subtlety over classical mathematics. To do so we will revisit the constructive notion of being finite and attack a second belief: that saying "a set is finite" is clear, concrete and self-explanatory. This thought is strengthened by the fact that a set in ZFC is either finite or infinite, tricking us to believe that all finite sets are so for the same reasons. We intend to demonstrate that it isn't so, that there are several "ways" of making a finite ZFC-set which are inherently distinct.

Indeed we will explore no less than four definitions of "finite" which are different (or conjecture to be so) in constructive mathematics even tough they are equivalent in ZFC. We will compare them, in terms of strength and closure properties, trying to hint what we gain or lose by choosing one rather than the other. Interestingly enough, none of these definitions contains a negation. This means in particular that the playful reader can create many more ways of being finite by inserting double negations in the definitions. However definitions without a negation are usually considerably more useful. Although we only explore these four notions,

Key words and phrases. Type theory, Constructive Logic, finite sets.

there are reasons to believe that there are many more, possibly infinitely many - all entirely positive.

Towards the end of the article we shall put these refinements of the idea of "finite" in motion by relating them with concrete, well-known, algorithms on graphs and finite automata. Indeed different ways of being finite mean different ways of exploiting finiteness in algorithms. We hope to convince the reader that constructive mathematics are a good tool to study algorithms precisely. Also that it is worth it not to restrict one's attention to a single, well chosen, notion of finite.

1.2. Notations.

We will use a set theory *à la* Bishop [1]. For a set A, we write $x \in A$ for x is a member of A. The set of functions from A to B is written $A \longrightarrow B$. For $f \in A \longrightarrow B$ and $g \in B \longrightarrow C$ we write $g \circ f$ for their compositions. For any set A, we write $\mathbf{1}$ the identity at A.

For two sets A and B, we define the cartesian product $A \times B$ together with the two projections π_1 and π_2. And the sum (or disjoint union) $A + B$ together with the two injections ι_1 and ι_2. We often omit the injection when they can be inferred from the context.

We consider as given the set of all propositions (where equality is equivalence) Ω[1]. We have therefore a set $A \longrightarrow \Omega$ of predicates over A. These predicates can be seen as subsets, we will often call them *parts*. Say P and Q are two parts of A. We write $a \in P$ instead of $P\ a$, effectively overloading this notation. We will write $P \cup Q$ and $P \cap Q$ for the point-wise disjunction and conjunction respectively. We note $\bot$ and $\top$ for the uniformly false and uniformly true part respectively. The set of parts of A can be ordered: we will write $P \subseteq Q$ to mean $\forall a.\ a \in P \to a \in Q$ - note that $P \subseteq Q \wedge Q \subseteq P$ is indeed equivalent to $P = Q$. Given a function $f \in A \longrightarrow B$ and a part $P \in A \longrightarrow \Omega$, we can define a part of B called the image of P through f and written simply $f\ P$ by $b \in f\ P\ =\ \exists a \in P. f\ a = b$. And given a function $f \in B \longrightarrow A$ we can define the inverse image of P through f as $b \in f^{-1}P = f\ b \in P$.

We will need the set $\mathbb{N}$ of natural numbers, and the sets $\mathbb{N}_{<n}$ of natural number smaller than n. We may write $\mathbb{B}$ for $\mathbb{N}_{<2}$ (the set containing precisely 0 and 1), $\mathbb{U}$ for $\mathbb{N}_{<1}$ and $\emptyset$ for $\mathbb{N}_{<0}$.

We write $\mathbb{L}_A$ for the set of lists of elements of A. The empty list is written $[]$, and list construction is noted $a :: l$. We can also write a list in extension $[a_1; a_2; \ldots; a_n]$. For $l \in \mathbb{L}_A$ we write $|l|$ for its length. Given a list $l \in \mathbb{L}_A$, we can define a part, $l^\Omega \in A \longrightarrow \Omega$ as:

- $[]^\Omega\ =\ \bot$
- $(a :: l)^\Omega\ x\ =\ x = a \vee x \in l^\Omega$

. Such a part is called enumerated. Also we say that a list has duplicates if it is of the shape $a :: l$ and $a \in l^\Omega$ or l has duplicates.

[1]The predicativists among the readers, might find that inappropriate. However this assumption is quite convenient and there is most likely a predicativist way to rewrite the associated parts of the article

We write $A^{\mathbb{N}}$ for the streams of elements of A. Given a stream $s \in A^{\mathbb{N}}$, a stream is given by its head $\mathsf{hd}\ s \in A$ and its tail $\mathsf{tl}\ s \in A^{\mathbb{N}}$. We write $a :: s$ for the stream such that $\mathsf{hd}\ (a :: s) = a$ and $\mathsf{tl}\ (a :: s) = s$. A list $l \in \mathbb{L}_A$ is said to be a prefix of a stream s if it can be obtained by "cutting" s at a certain position[2]. Given a position k, we can also *remove the k-th prefix from s* - written $s|_k \in A^{\mathbb{N}}$ - defined inductively as follows:

- $s|_0 = s$
- $s|_{Sk} = \mathsf{tl}\ s|_k$

We sometimes use the notation $s_k = \mathsf{hd}\ s|_k$ for the k-th element of s. A stream s is said to have duplicates if it has a prefix with duplicates.

Finally we call a property over sets a mapping P, mapping each set to a proposition, so that too isomorphic sets are mapped to equivalent propositions. We write $A\ \varepsilon\ P$ for PA. And we call a set A decidable if there is a function $e \in A \longrightarrow A \longrightarrow \mathbb{B}$ so that $e\ x\ y = 1$ if and only if $x = y$. For instance, $\mathbb{N}$ and all the $\mathbb{N}_{<n}$ are decidable.

2. 'Finite'? Did you say 'finite'?

2.1. Enumerated sets.

Definition
A set A will be said to be *enumerated*, or simply finite, noted $A\ \varepsilon\ \mathcal{F}$ when there is a list of all its elements. We call such a list an *enumeration* of A.

This is the most common notion used to characterise finiteness in constructive mathematics. Although it is arguably the notion which deserves the name "finite" the best, going back to its etymology: "complete" or "finished", we will avoid calling enumerated sets finite to prevent confusion. Likewise "finiteness" will refer to any of the notion presented in this article.

Notice that it is, in particular, decidable for enumerated set whether they are inhabited.

Property
Let $A\ \varepsilon\ \mathcal{F}$ and $P \in A \longrightarrow \mathbb{B}$, then $\{x \in A \mid P\ x\}$ is enumerated.
This is not true however if we take P in $A \longrightarrow \Omega$. In particular, if $f \in B \longrightarrow A$ is injective, we cannot conclude that $B\ \varepsilon\ \mathcal{F}$.

> The decidable case is clear: just filter out those elements of the enumeration of A where P is 0 and you get an enumeration of $\{x \in A \mid P\ x\}$. For the general case, let us consider the one element set $\mathbb{U}$, it can be enumerated by the list $[0]$. Let A be an arbitrary proposition, and P be

[2]Formally, it is defined inductively on l:

- $[]$ is a prefix of any stream s
- $a :: l$ is a prefix of s if $a = \mathsf{hd}\ s$ and l is a prefix of $\mathsf{tl}\ s$.

the constant predicate $\lambda x, A \in \mathbb{U} \longrightarrow \Omega$. If $\{x \in \mathbb{U} \mid P\ x\}$ is enumerated then by checking whether it is inhabited, we can decide A.

Property
Let $A\ \varepsilon\ \mathcal{F}$ and $f \in A \longrightarrow B$ a surjective function. Then $B\ \varepsilon\ \mathcal{F}$.

Indeed, if l is an enumeration of A, then fl is an enumeration of B (as f is surjective).

Property
If $A\ \varepsilon\ \mathcal{F}$ and $B\ \varepsilon\ \mathcal{F}$, then $A \times B\ \varepsilon\ \mathcal{F}$ and $A + B\ \varepsilon\ \mathcal{F}$.

If l and v are respective enumerations of A and B, we can build an enumeration of $A \times B$ by taking the list of all (s, t) with s in l and t in v, it is obtained by binary list comprehension. An enumeration of $A + B$ is obtained by appending l and v.

2.2. Bounded size sets.

Definition
A set A is said to have *bounded size*, noted $A\ \varepsilon\ \mathcal{B}$ if there is a natural number N, such that every list $l \in \mathbb{L}_A$ with $|l| > N$ has duplicates. We call such an N a bound on the size of A.

This is an other common notion. It takes quite a step from enumeration, as it doesn't give a way to choose an element in A. On the plus side, it is quite robust.

Property
Let $A\ \varepsilon\ \mathcal{B}$ and $f \in B \longrightarrow A$ an injective function , then $B\ \varepsilon\ \mathcal{B}$.

Indeed, for $l \in \mathbb{L}_B$, if $f\ l$ has duplicates, then, as f is injective, l also has duplicates. Therefore the bound on the size of A is a bound on the size of B.

Property
Any enumerated set has bounded size. However the converse is not true.

Let $A\ \varepsilon\ \mathcal{F}$ and l an enumeration of A. Then for any list $v \in \mathbb{L}_A$ of length greater than that of l, by the *pigeonhole principle*, there are duplicates. Therefore $|l|$ is a bound on the size of A. Additionally, since $\mathcal{B}$, unlike $\mathcal{F}$, is stable under arbitrary subsets, we cannot find an enumeration of an arbitrary bounded size set.

Property
Let $A\ \varepsilon\ \mathcal{B}$ and $f \in A \longrightarrow B$ surjective, then $B\ \varepsilon\ \mathcal{B}$.

Let N be a bound on the size of A and $l \in \mathbb{L}_B$ such that $|l| > N$. Since f is surjective, there is a list v such that $l = f\ v$. We have $|v| > N$, thus v has duplicates, therefore $l = f\ v$ has duplicates. We can conclude that N is a bound on the size of A.

Property
Let $A\ \varepsilon\ \mathcal{B}$ and $B\ \varepsilon\ \mathcal{B}$. $A \times B\ \varepsilon\ \mathcal{B}$ and $A + B\ \varepsilon\ \mathcal{B}$.

Let N and M be, respectively, bounds on the size of A and B.

We will call two elements of $A \times B$ (or a larger product) A-equal (resp. B-equal) if their first (resp. second) projections are equal. We also have the corresponding notion of A-duplicate (resp. B-duplicates) for lists of $\mathbb{L}_{A\times B}$ (and lists over larger products). $L \in \mathbb{L}_{A\times B\times\mathbb{L}_{\mathbb{N}}}$. If L has length longer than N, we can find a pair of A-duplicates in it. say (s_1, t_1, l_1) and (s_2, t_2, l_2). We can then remove them from the list and replace them by $(s_1, t_1, l_1{+\!\!+}l_2)$ (where $l_1{+\!\!+}l_2$ denotes the concatenation of both lists). Now let $l \in \mathbb{L}_{A\times B}$ such that $|l| > NM$. And set L_0 to be the list obtained from l by adjoining the singleton list of its position to each element. The following two properties are verified on L_0 and preserved by the above operation:

- The concatenation of the contained list is a permutation of $\mathbb{N}_{<|l|}$
- If the position i is in the sublist accompanying the element (s, t), then the i-th element of l is A-equal to (s, t)

Starting with L_0 we can iterate the operation until we are reduced to a list of size N (or less, as NM can be 0). We extract the lists of positions, giving at most N lists of total size greater than NM, therefore one of them, we will call it p, has length greater than M. Let l' the sublist of l corresponding to the positions in p, since $|l'| > M$, it has a pair of B-equal elements. And since all elements of l' are A-equal, these two elements are in fact equal. We conclude that NM is a bound on the size of $A \times B$.

For $A+B$ it is more straightforward, let $l \in \mathbb{L}_{A+B}$ with $|l| > N+M$. Let l_A be the sublist of l of the elements of A and l_B that of the elements of B. Either $|l_A| > N$ or $|l_B| > M$, in both case we find duplicates in A. Thus $N + M$ is a bound on the size of $A + B$.

2.3. Noetherian sets.

Definition
Let A be a set, for a list $l \in \mathbb{L}_A$ we define the following property inductively:

- If l has duplicates then $A\ \varepsilon\ \mathcal{N}_l$
- If for all $a \in A$, $A\ \varepsilon\ \mathcal{N}_{a::l}$, then $A\ \varepsilon\ \mathcal{N}_l$

A is said to be noetherian, written $A\ \varepsilon\ \mathcal{N}$ if $\forall l.\ A\ \varepsilon\ \mathcal{N}_l$ or, equivalently, if $A\ \varepsilon\ \mathcal{N}_{[]}$.

This notion seems to be new. It provides an induction principle over lists. Which will prove useful. It is also reminiscent of the definition of noetherian rings - hence the name. To emphasize this, we could rephrase the definition as "any ascending chain of enumerated subsets has two consecutive equal terms" (in classical mathematics we would say that ascending chains eventually stabilise, but constructiveness binds us to more finitary definitions when possible).

Building on this thought we can give an alternate definition of noetherian sets without referring to *enumerated* sets.

Definition
Let A be a set, and $P \in A \longrightarrow \Omega$, we define inductively $P \in \mathcal{N}^A$:

- $P \in \mathcal{N}^A$ if for all $a \in A$, $a \in P$ or $P \cup [a]^\Omega \in \mathcal{N}^A$

Property
If $P \subseteq Q$, and $P \in \mathcal{N}^A$ then $Q \in \mathcal{N}^A$

By induction on the proof of $P \in \mathcal{N}^A$:

- If $a \in P$ then $a \in Q$
- If $P \cup [a]^\Omega$ then, by induction hypothesis, $Q \cup [a]^\Omega$ is finite (as $P \cup [a]^\Omega \subseteq Q \cup [a]^\Omega$)

Property
A set A is noetherian if and only if $P \in \mathcal{N}^A$ for all P, or equivalently $\bot \in \mathcal{N}^A$.

We will first prove that $A \ \varepsilon \ \mathcal{N}_l$ if and only if $l^\Omega \in \mathcal{N}^A$ or l has duplicates. Since [] doesn't have duplicates, the result follows.

- If $A \ \varepsilon \ \mathcal{N}_l$, by induction:
  - If l has duplicates then we have our result directly
  - If, for all $a \in A$, $A \ \varepsilon \ \mathcal{N}_{a::l}$ then, by induction hypothesis, either
    - $a :: l^\Omega \in \mathcal{N}^A$
    - $a :: l$ has duplicates, that is $a \in l^\Omega$

    Hence our result.
- For the other direction we have two cases:
  - l has duplicates, therefore $A \ \varepsilon \ \mathcal{N}_l$
  - $l^\Omega \in \mathcal{N}^A$, then we prove the result by induction. For any $a \in A$, either:
    - $a \in l^\Omega$, thus $a :: l$ has duplicates and $A \ \varepsilon \ \mathcal{N}_l$
    - $a :: l^\Omega = l^\Omega \cup [a]^\Omega \in \mathcal{N}^A$, which, by induction hypothesis, implies that $A \ \varepsilon \ a :: l$

    Therefore $A \ \varepsilon \ l$.

Property
Any set with bounded size (hence any enumerated set) is noetherian. However, the converse is not necessarily true.

It is straightforward to prove, by induction on $N + 1 - |l|$, that if N is a bound on the size of A, then $A\ \varepsilon\ \mathcal{N}_l$.

The following example, showing that the converse does not hold has been suggested by F. Richman. Consider the following family of subsets of $\mathbb{N}$: $a_0 = [0, 0]$, $a_1 = [1, 2]$, $a_2 = [3, 5]$ $a_3 = [6, 9]$ and so on. We will use the following properties of this sequence: a_i has exactly $i + 1$ different elements and for a given natural number k, there is a unique νk such that $k \in a_{\nu k}$. Now we need an arbitrary stream $s \in \mathbb{B}^{\mathbb{N}}$. We construct the set $X_s = \{k \in \mathbb{N} \mid s_{\nu k} = 1 \wedge \forall j < \nu k.\ s_j = 0\}$. so that X_s has the same elements as one of the a_i or is empty. Giving a bound N on the size of X_s gives a bound on the position of the first 1 in s. In particular s is identically 0 if and only if for all $i < N$, $s_i = 0$. Which is obviously decidable. Therefore if I can give a bound on the size of X_s, I can decide whether s is identically 0, which isn't plausible.

On the other hand, X_s is noetherian. Indeed given an element k of X_s, we can get a bound on the size of X_s, namely νk.

Property
Let $A\ \varepsilon\ \mathcal{N}$ and $f \in B \longrightarrow A$ an injective function, then $B\ \varepsilon\ \mathcal{N}$.

Let us prove, by induction that if $P \in \mathcal{N}^A$ then $f^{-1}P \in \mathcal{N}^B$, for any $b \in B$:

- If $f\ b \in P$, then $b \in f^{-1}P$
- If $P \cup [f\ b]^{\Omega} \in \mathcal{N}^A$ then, by induction hypothesis, $f^{-1}(P \cup [f\ b]^{\Omega})$ which, since f is injective, equals $f^{-1}P \cup [b]^{\Omega}$ is $\mathcal{N}^B$

Property
Let $A\ \varepsilon\ \mathcal{N}$ and $f \in A \longrightarrow B$ a surjective function, then $B\ \varepsilon\ \mathcal{N}$.

We shall prove, by induction, that if $A\ \varepsilon\ \mathcal{N}_l$ then $B\ \varepsilon\ \mathcal{N}_{fl}$.

- If l has duplicates, then $f\ l$ too, hence $B\ \varepsilon\ \mathcal{N}_{fl}$
- We have $A\ \varepsilon\ \mathcal{N}_l$, let $b \in B$, as f is surjective there is an $a \in A$ such that $b = f\ a$. We have $A\ \varepsilon\ \mathcal{N}_{a::l}$, then, by induction hypothesis, $B\ \varepsilon\ \mathcal{N}_{b::(fl)}$. Hence the result.

Property
If $P \in \mathcal{N}^A$ and $Q \in \mathcal{N}^B$ then $P \cap Q \in \mathcal{N}^A$.

By simultaneous induction, for any $a \in A$:

- If $a \in P$ and $a \in Q$, then $a \in P \cap Q$

- If $P \cup [a]^\Omega \in \mathcal{N}^A$ and $a \in Q$ then, by induction hypothesis, $(P \cap Q) \cup [a]^\Omega = (P \cup [a]^\Omega) \cap Q \in \mathcal{N}^A$
- Idem if $a \in P$ and $Q \cup [a]^\Omega \in \mathcal{N}^A$
- If $P \cup [a]^\Omega \in \mathcal{N}^A$ and $Q \cup [a]^\Omega \in \mathcal{N}^A$, then, by induction hypothesis, $(P \cap Q) \cup [a]^\Omega = (P \cup [a]^\Omega) \cap (Q \cup [a]^\Omega) \in \mathcal{N}^A$

Property
If $A \varepsilon \mathcal{N}$ and $B \varepsilon \mathcal{N}$, then $A \times B \varepsilon \mathcal{N}$ and $A + B \varepsilon \mathcal{N}$.

Let $P \in A \longrightarrow \Omega$ and $Q \in B \longrightarrow \Omega$, we shall write $P|Q \in A \times B \longrightarrow \Omega$ for the predicate defined by $(a, b) \in P|Q = a \in P \vee b \in Q$. Now we will prove that for all $P \in \mathcal{N}^A$, $Q \in \mathcal{N}^B$, $P|Q \in \mathcal{N}^{A\times B}$. As $\bot|\bot = \bot$, the result follows. By simultaneous induction, for any $(a, b) \in A \times B$:

- If either $a \in P$ or $b \in Q$ then, $(a, b) \in P|Q$
- If $P \cup [a]^\Omega \in \mathcal{N}^A$ and $Q \cup [b]^\Omega \in \mathcal{N}^B$, then by induction hypothesis both $P \cup [a]^\Omega|Q$ and $P|Q \cup [b]^\Omega$ are $\mathcal{N}^{A\times B}$, therefore $P \cup [a]^\Omega|Q \cap P|Q \cup [b]^\Omega = P|Q \cup [(a, b)]^\Omega \in \mathcal{N}^{A\times B}$

For sums, we can use a similar - slightly simpler - proof. For $P \in A \longrightarrow \Omega$, $Q \in B \longrightarrow \Omega$, let $P+Q \in A+B \longrightarrow \Omega$ such that $\iota_1 a \in P+Q = a \in P$ and $\iota_2 b \in P + Q = b \in P$.

We shall prove that when $P \in \mathcal{N}^A$ and $Q \in \mathcal{N}^B$ then $P+Q \in \mathcal{N}^{A+B}$. Again, $\bot + \bot = \bot$, from which we may conclude. By induction, forall $x \in A + B$:

- If $x = \iota_1 a$, then either
  - $a \in P$, then $a \in P + Q$
  - $P \cup [a]^\Omega \in \mathcal{N}_A$, and, by induction hypothesis, $P + Q \cup [x]^\Omega = P \cup [a]^\Omega + Q \in \mathcal{N}^{A+B}$
- Symmetrically, if $x = \iota_2 b$.

Property
If $A \varepsilon \mathcal{N}$ is decidable then, the set $\{l^\Omega \mid l \in \mathbb{L}_A\}$ of enumerated parts of A is also noetherian.

We prove the property for the set A^+ of non-empty enumerated parts first. then as $\{l^\Omega \mid l \in \mathbb{L}_A\}$ is isomorphic to $A^+ + \mathbb{U}$, we conclude.

For $P \in A \longrightarrow \Omega$, let $P^+ \in A^+ \longrightarrow \Omega$ defined by $p \in P^+$ if $\exists a \in P.\ a \in p$. As $\bot^+ = \bot$, it suffices to prove that when $P \in \mathcal{N}^A$, we have $P^+ \in \mathcal{N}^{A^+}$.

Before going on, we notice that, as A is decidable, the part $\overline{a}$ defined by $p \in \overline{a}$ when $a \notin p$ is $\mathcal{N}^{A^+}$. We can procede by induction:

- We know that for all $a \in A$, $a \in P$ or $(P \cup [a]^\Omega)^+ \in \mathcal{N}^{A^+}$, We want to prove that $P^+ \in \mathcal{N}^{A^+}$. It suffices to show that for all $(a :: l)^\Omega \in A^+$, $(a :: l)^\Omega \in P^+$ or $P^+ \cup [(a :: l)^\Omega]^\Omega \in \mathcal{N}^{A^+}$.

- If $a \in P$ then $(a :: l)^{\Omega} \in P^{+}$
- If $(P \cup [a]^{\Omega})^{+} \in \mathcal{N}^{A^{+}}$, then $(P \cup [a]^{\Omega})^{+} \cap \overline{a} \subseteq P^{+} \cup [(a :: l)^{\Omega}]^{\Omega}$. Since the former is $\mathcal{N}^{A^{+}}$ the latter is so as well.

In the above proof the hypothesis of decidability is used critically proving $\overline{a}$ to be $\mathcal{N}^{A^{+}}$. But we could loosen the hypothesis a bit and consider an arbitrary noetherian sets and only those parts that are enumerated by duplicate-free lists (then we would cut the considered set in three part: $\mathbb{U}$, A, and the set of those parts which are enumerated by duplicate-free lists of size at least 2). It remains an open question whether we could simply drop the decidability hypothesis.

2.4. Streamless sets.

Definition
A set A is said streamless, noted $A \; \varepsilon \; \overline{\mathcal{S}}$, if every stream $s \in A^{\mathbb{N}}$ has duplicates. Equivalently, for any $s \in A^{\mathbb{N}}$ there are two positions $i < j$ such that $s_i = s_j$.

This notion is fairly close to that of classical mathematics wherein a set is finite if there is no injection from $\mathbb{N}$ to it; though formulated in a more positive fashion.

We can show easily that it is a weaker notion than that of being noetherian:

Property
For any $A \; \varepsilon \; \mathcal{N}$, we have that $A \; \varepsilon \; \overline{\mathcal{S}}$.

The easiest path is to define, for a list $l \in \mathbb{L}_A$ and a stream $s \in A^{\mathbb{N}}$, the "reversed concatenation" $l \star s$:

- $[] \star s \;=\; s$
- $(a :: l) \star s \;=\; l \star (a :: s)$

It suffices that if $A \; \varepsilon \; \mathcal{N}_l$, then for all stream $s \in A^{\mathbb{N}}$, $l \star s$ has duplicates.

This is proven by a straightforward induction.

We conjecture that there is a set A which is streamless but cannot be proven noetherian. The main evidence to support this conjecture is that the statement $\forall \mathrm{A}.\; A \; \varepsilon \; \overline{\mathcal{S}} \to A \; \varepsilon \; \mathcal{N}$ is an instance of the well studied principle of bar-induction [2] where "has duplicates" plays the role of the bar. Bar-induction is well-known *not* to be valid in intuitionistic mathematics [3].

Property
If $A \; \varepsilon \; \overline{\mathcal{S}}$ then $A + \mathbb{U} \; \varepsilon \; \overline{\mathcal{S}}$.

Let $s \in (A + \mathbb{U})^{\mathbb{N}}$, if we peek at the two first elements:

- Either they are both 0 and we're done
- Either one of them is a member of A, which we will call a_0. We can then construct a stream $s' \in A^{\mathbb{N}}$ by taking s and replacing all of its 0-s by a_0.
  As s' is a stream in A, we get two position $i < j$ such that ${s'}_i = {s'}_j$. Likewise, we have two position $i' < j'$ such that $\left(s'|_{j+1}\right)_{i'} =$

$(s'|_{j+1})_{j'}$. By setting $k = j + 1 + i'$ and $l = j + 1 + j'$. We have $i < j < k < l$ such that ${s'}_i = {s'}_j$ and ${s'}_k = {s'}_l$. Now consider s_i, s_j, s_k and s_l. Either two of them are 0 and we're done. Either at least one of the latter equations works with s instead of s', which proves the result.

Property
If $A \varepsilon \overline{\mathcal{S}}$ and $B \varepsilon \overline{\mathcal{S}}$ then $A + B \varepsilon \overline{\mathcal{S}}$.

Given a stream $s \in A + B^{\mathbb{N}}$, we shall write $s^A \in A + \mathbb{U}^{\mathbb{N}}$ for the stream obtained from s by replacing all $\iota_2 b$ by 0. And likewise we define $s^B \in B + \mathbb{U}^{\mathbb{N}}$.

Now, we need to prove for an arbitrary stream $s \in A + B^{\mathbb{N}}$ that it has two positions $i < j$ such that $s_i = s_j$. Let us first define the stream $s' \in A + B^{\mathbb{N}}$:

- To define $\mathsf{hd}\ s'$, we consider s^A, it has two positions $i < j$ such that ${s^A}_i = {s^A}_j$. We set $\mathsf{hd}\ s'$ to be s_i
- To define $\mathsf{tl}\ s'$, we continue "corecursively", acting on $s^A|_{j+1}$

s' has the property that each of its positions k correspond to two positions $Lk < Rk$ in s such that:

- $Rk < L(k + 1)$
- If ${s'}_k \in B$ then $s_{Lk} = {s'}_k$
- If ${s'}_k \in A$ then $s_{Lk} = s_{Rk}$

Then ${s'}^B$ has two positions $i < j$ such that ${s'^B}_i = {s'^B}_j$. Either both are in B or both are 0; in either case the above properties suffice to conclude.

It is an open question, however, whether the cartesian product of two streamless sets is itself streamless. A partial answer can be found in [4].

3. Algorithms

3.1. Concrete parts.

Before going on we need a slight sophistication of the notion of part. Parts are indeed a powerful specification tool but a poor algorithmic object. We at least need decidable predicates for algorithmic purposes, but it is often not enough: we will need to list all the elements in the part.

Definition
We shall call a set of *concrete parts* of a set A a set Γ endowed with:

- a member function $(\in) \in \Gamma \longrightarrow A \longrightarrow \mathbb{B}$
- a list function $! \in \Gamma \longrightarrow \mathbb{L}_A$
- a union function $(\cup) \in \Gamma \longrightarrow \Gamma \longrightarrow \Gamma$
- an intersection function $(\cap) \in \Gamma \longrightarrow \Gamma \longrightarrow \Gamma$
- an empty element $\bot \in \Gamma$
- an addition function $(\cdot) \in A \longrightarrow \Gamma \longrightarrow \Gamma$

such that:

- For all $\gamma \in \Gamma$, $(!\gamma)^{\Omega} = (\in\gamma)$
- $\in(\gamma_1 \cup \gamma_2) = (\in\gamma_1) \cup (\in\gamma_2)$
- $\in(\gamma_1 \cap \gamma_2) = (\in\gamma_1) \cap (\in\gamma_2)$
- $(\in\bot) = \bot$
- $(\in(\mathrm{a}\cdot\gamma)) = (\in\gamma) \cup [a]^{\Omega}$

Where we abuse notations in assimilating decidable predicates to parts.

The above definition simply makes precise the idea that "a concrete part is a decidable part that can be listed". We set our notation specifically so that we can manipulate concrete parts just as parts. Also we will, from now on, omit the reference to a particular type of concrete parts and speak of concrete parts, remembering implicitly that they need to be of the same type to be combined.

Notice that, for a set A, having a type of concrete parts implies that A is decidable. Indeed $a = b$ if and only if $a \in (b \cdot \bot)$. This remarks shows that a concrete part is essentially the same as an enumerated part of a decidable set. Defining them as such would cause trouble when it comes to complexity, as lists are a rather bad implementation of concrete parts. Typical efficient implementations are hash tables and balanced binary search tree. Both require extra properties, but we won't need to choose a particular implementation.

3.2. Reachability.

In this subsection we will focus our interest on the standard reachability algorithm for oriented graphs. Let us fix a graph, by giving a set $\mathcal{V}$ of vertices, and for each vertex v a concrete part $\mathcal{N}_v$ of successors. Given a vertex $v \in \mathcal{V}$ in a graph, we want to compute the concrete part $\mathcal{R}_v \in \mathcal{V} \longrightarrow \Omega$ of vertices u such that there is a path from v to u.

The algorithm goes as follows: set the the current vertex to be v and no vertex to be marked but v, then

- Let v be the current vertex. For each $u \in \mathcal{N}_v$, if u is unmarked then mark it, and push it on top of a pile, else do nothing.
- Pop the top of the pile, and set it as the current vertex
- Start over until the pile is empty

At the end of the algorithm the concrete part of marked vertices is precisely $\mathcal{R}_v$.

Now, for this to terminate, $\mathcal{V}$ obviously needs to verify some finiteness property. We're left with the question of which of those we know fits best. The intuitive argument goes as follows: "as $\mathcal{V}$ is finite, and we're feeding a part of it with vertices that weren't there before, this has to stop somewhere". This sets us in the realm of noetherian sets. It is possible to make the argument precise using the following inductive definition:

- A part $P \in A \longrightarrow \Omega$ is negatively close to the top of A if for all $a \notin P$, $P \cup [a]^{\Omega}$ is negatively close to the top of A.

As $\mathcal{V}$ is decidable, we have that P is negatively close to the top of $\mathcal{V}$ if and only if $P \in \mathcal{N}^A$. By lexicographic induction on the proof that P is negatively close

to the top of $\mathcal{V}$ and on the size of the pile, we can conclude that the algorithm terminates.

Now would being streamless be a sufficient condition for $\mathcal{V}$ to let the procedure terminate? It would seem so, even though the argument is a bit less conventional: we build a stream $s \in (\mathcal{V} + \mathbb{U})^{\mathbb{N}}$ by following the algorithm and "outputing" to s every u when it is being marked, also when the pile is empty we need to "output" 0 to s. This gives a productive stream regardless of whether the procedure terminates. Also at two positions $i < j$ either $s_i \neq s_j$ or $s_i = s_j = 0$. Since we chose $\mathcal{V}$ to be streamless, we get two positions $i < j$ such that $s_i = s_j$ thus $s_i = 0$. Therefore i is a bound on the number of step to reach the empty pile for this particular v.

3.3. Automata and determinisation.

An automaton is little more than a graph. It has a set of states $\mathcal{V}$, a concrete part $\mathcal{I}$ of which are initial, a decidable part $\mathcal{F}$ of final states[3]. For each state $v \in \mathcal{V}$ and symbol a in the (enumerated) alphabet, there is a concrete part $\mathcal{N}_{v,a}$ of successors [4].

Reachability is an important problem in finite automata, as it allows, for instance, to decide whether the recognised language is empty. But this doesn't constrain us much since we established, in previous section, that $\mathcal{V}$ need only be streamless for this problem to be solved. On the other hand there is a need of taking the cartesian product of the sets of vertices of two automata (to recognise the intersection of their languages, for instance). We could only conjecture that it was possible for streamless sets - it works for decidable noetherian sets though.

Let us have a glimpse in the direction of automaton determinisation. We're given an automaton and we want an equivalent automaton which is deterministic - that is where all the $\mathcal{N}_{v,a}$ contain at most one element and $\mathcal{I}$ contains exactly one element. This is a well known practice and goes as follows:

- The set of states of the determinised automaton is a set Γ of concrete parts of $\mathcal{V}$
- The initial state is $\mathcal{I}$
- A state $\gamma \in \Gamma$ is final if one of its element is in $\mathcal{F}$
- The concrete part $\mathcal{N}_{\gamma,a}$ is the union of all the $\mathcal{N}_{u,a}$ for $u \in \gamma$

This works as long as the set of Γ of states of the determinised automaton is sufficiently finite - that is, at least as finite as what is required for an automaton. Γ is isomorphic to the set of enumerated parts of $\mathcal{V}$, the set of enumerated parts of a decidable noetherianness set is noetherian. We have no evidence that this could hold for streamless sets. It is therefore a safe bet to assume that an automaton needs to have a noetherian set of states so that it can be determinised.

[3]It is often useful to have actually a concrete part of final states, but this would only make things obfuscated in our case

[4]We consider non-deterministic automata without ϵ-transitions

4. Conclusion

4.1. Related works.

This article is strongly related to [4] via Sections 2.3 and 2.4. To make this statement more precise, being streamless is the same as what Richman and Stolzenberg call **2**-good (a set can be seen as a preorder where $x \leq y$ means $x = y$). Being noetherian, for a decidable set, is the same as being a well quasi-order in Richman and Stolzenberg sense.

Actually their article focuses on preorders, but it doesn't seem essential, we could generalize "goodness" and "wellness" to arbitrary relation. Our definition of noetherian is stronger than their definition of being well (in their common case). And that is because they use a negation. Our Section 2.3, therefore, suggests a new definition of well which goes, roughly, as follows: If A is a set and $R \in A \longrightarrow A \longrightarrow \Omega$ a relation, we define inductively for a part $P \in A \longrightarrow \Omega$ the following property:

- R is well from P on if for all $a \in A$ either $R\ a \subseteq P$ or R is well from $P \cup (R\ a)$ on.

And R is well if it's well from $\bot$ on.

This definition coincides with noetherianty when R is chosen to be equality. It has the advantage over the notion in [4] to avoid the need of a decidability hypothesis on many lemmas.

Richman and Stolzenberg also propose a family of properties which are in-between **2**-good and well. This would require a careful study as some of them are equivalent in the equality case; it might reveal that there are indeed infinitely many finiteness properties between noetherian and streamless. On top of this family lies a notion that could be worth special attention, which Richman and Stolzenberg call *bar-good*.

4.2. A few last words.

Here ends our little tour in constructive mathematics and finiteness. We hope to have convinced the reader of their relevance, at least as far the study of algorithms is concerned.

We are left with a few questions. Is streamlessness stable by cartesian product? Is the set of enumerated part of a noetherian set also noetherian even if the set isn't decidable ? What about when the set is only streamless ? It is also conjectured that we can find a set which is streamless but cannot be proven to be noetherian. And there is quite probably many other questions to find.

References

[1] E. Bishop. *Foundations of Constructive Analysis*. McGraw-Hill, New York, 1967.
[2] W. A. Howard, G. Kreisel. Transfinite induction and bar induction of types zero and one, and the role of continuity in intuitionistic analysis. *J. Symb. Log.* **31**(3), 325–358, 1966.
[3] S. C. Kleene, R. E. Vesley. *The Foundations of Intuitionistic Mathematics*. North–Holland, 1965.
[4] F. Richman, G. Stolzenberg. Well quasi-ordered sets. *Avdanced Mathematics*, pp. 145–153, 1993.

Chalmers Tekniska Högskola, Göteborg, Sweden
E-mail address: coquand@cs.chalmers.se

Equipe TypiCal, LIX, Ecole Polytechnique, France
E-mail address: Arnaud.Spiwack@lix.polytechnique.fr