An Induction Principle over Real Numbers

Assia Mahboubi

AN INDUCTION PRINCIPLE OVER REAL NUMBERS

ASSIA MAHBOUBI

Abstract. We give a constructive proof of the open induction principle on real numbers, using bar induction and *enumerative* open sets. We comment the algorithmic content of this result.

1. Introduction

We call *open induction over real numbers* an elementary lemma of real analysis which states that an open set of the closed interval $[0,1]$ satisfying an inductive property in fact covers entirely the space $[0,1]$.

Theorem 1.1 (Open Induction). *Let A be an open set of the closed interval $[0,1]$, satisfying the following property (A is said to be inductive) :*

$$\forall x \in [0,1],\ (\forall y \in [0,x[, y \in A) \Rightarrow x \in A$$

Then $A = [0,1]$.

Property $(*)$ is said to be *inductive* because it allows to prove that an element of $[0,1]$ is in the open set A as soon as all its predecessors for the order over $\mathbb{R}$ are themselves in A. In the more general context of universal algebra, this property was first formulated as an induction principle and studied from a classical viewpoint by J.-C. Raoult in [3]. In this paper we propose a constructive proof of this result for real numbers equipped with the usual topology and hence explicit its computational content. The classical proof of Theorem 1 is quite straightforward. On the other hand a constructive one requires the use of an extra non-classical axiom.

The proof we give below follows closely that of Th. Coquand [1] for the Cantor dyadic set. Both rely on an encoding of the problem in finite words of natural numbers and build the proof that the space considered is included in the open set piecewise, using a function which gives the current state of the proof and a predicate which builds the certificates of inclusion. They are both proofs by *bar induction.* All the arguments used in this development hold classically although the whole proof becomes then unnecessarily complicated from a classical point of view.

Nevertheless, there are some strong differences between the proofs. The proof presented below deals directly with axiomatized real numbers and the bar property strongly relies on the completeness of $\mathbb{R}$ through the use of the shrinking nested intervals property, whereas Th. Coquand's proof [1] avoids the need of a completeness argument by working over $\{0,1\}^{\mathbb{N}}$. Moreover the enumerative open sets we work with seem not to be the same as the ones provided by the canonical surjection of the Cantor set. Our definition is in fact the one used by W. Veldman in [7] where he also gives a proof of the open induction principle. This latter proof relies on a strong version of the fan theorem instead of bar induction, and deals with real numbers as Cauchy fundamental sequences (see [7] and [4]) whereas that presented here stands in any *real numbers structure* (see [2]), whatever model has been chosen.

2. Notation and definitions

2.1. **About bar induction.** In all what follows, $\mathbb{N}$ denotes the set of natural numbers. Thus $\mathbb{N}^*$ refers to the set of finite words over $\mathbb{N}$ and $\mathbb{N}^\mathbb{N}$ to the set of infinite sequences over $\mathbb{N}$.
ε denotes the empty word.
If α is in $\mathbb{N}^\mathbb{N}$ and x in $\mathbb{N}$, then $\alpha(x)$ stands for the finite prefix of α of length x, $\alpha_1 \dots \alpha_x$.
If u is a finite word in $\mathbb{N}^*$ and a a natural number, $u \bullet a$ stands for the concatenation of u and the word a of length 1.
If u is a finite word, $|u|$ denotes its length.

$\mathbb{N}^*$ can be seen as the *universal tree*, whose infinite branches are the set $\mathbb{N}^\mathbb{N}$. A finite word of $\mathbb{N}^*$ is then the position of the associated node in T. Then a bar is a set of nodes which cannot be avoided when constructing and infinite path in T starting from the root.

Definition 2.1 (Bar). Let X be a predicate over $\mathbb{N}^*$. X is called a *bar* if every infinite sequence of natural numbers has a finite prefix satisfying X:

$$\forall \alpha \in \mathbb{N}^\mathbb{N}\ \exists k \in \mathbb{N}\ X(\alpha(k))$$

The axiom of bar induction is a non classical axiom which can be stated in several ways (see for instance [4] and [6] for a complete survey of the different versions and their relations). Its main consequence is the fan theorem (intuitionistic version of König's lemma) and both are largely involved in the intuitionistic reconstruction of real analysis (see [4] and [5]). This principle can be seen as an induction principle over the universal tree T. From this point of view, given Y a property which holds on a bar and is transmitted upwards (Y is progressive), the bar induction principle allows to conclude that Y holds for the root ε.

We will here use the axiom of decidable bar induction, which can be expressed using the following inference rule:

Axiom 1 (Bar induction). *Let X and Y be two predicates over $\mathbb{N}^*$.*

$$\frac{\begin{array}{ll} \forall u \in \mathbb{N}^*\ [X(u) \Rightarrow Y(u)] & \textit{X implies Y} \\ \forall \alpha \in \mathbb{N}^\mathbb{N}\ \exists k \mid X(\alpha(k)) & \textit{X is a bar} \\ \forall u\ X(u) \vee \neg X(u) & \textit{X is decidable} \\ \forall u \in \mathbb{N}^*\ [\forall a \in \mathbb{N}\ Y(u \bullet a)] \Rightarrow Y(u) & \textit{Y is progressive} \end{array}}{Y(\varepsilon)}$$

2.2. **Real numbers and enumerative open sets.** The set $\mathbb{R}$ of real numbers is axiomatically defined as a Cauchy complete Archimedean ordered field. The set $\mathbb{Q}$ of rational numbers is a countable, dense subset of $\mathbb{R}$. Equality between rational numbers is decidable.

Definition 2.2 (Dyadic numbers). A dyadic number is a rational number which can be written as a quotient whose denominator is a power of 2.

The set of dyadic numbers belonging to $[0,1]$ is also called the dyadic Cantor set and denoted by $\{0,1\}^{\mathbb{N}}$

We now define the open sets we are going to work with. The choice of such a definition is crucial as it is closely linked to the termination of the computation of the proof. For example, defining real open sets by giving for each of its element the ray of an open ball included in the open set seems not to be convenient to the kind of argument we use.

Definition 2.3 (Enumerative open sets)**.** An enumerative open set A of $\mathbb{R}$ is a countable union of open intervals with rational bounds given by a mapping :

$$\begin{array}{rccc} g: & \mathbb{N} & \rightarrow & \mathbb{Q}\times\mathbb{Q} \\ & n & \mapsto & (\alpha_n,\beta_n) \end{array}$$

Namely $A = \bigcup_{n\in\mathbb{N}}]\alpha_n, \beta_n[$

These enumerative open sets need not be decidable. Yet the inclusion of an interval with rational bounds in a finite union of open intervals with rational bound is obviously decidable, which will be sufficient for the bar involved to be decidable (see lemma 3.2 below).

If E is a subset of $\mathbb{R}$, an open set of E is the intersection of an enumerative open set with E.

Finally if I is the closed real interval $[x,y]$, then we denote by I_l the closed interval $[x, \frac{x+y}{2}]$ and by I_r the closed interval $[\frac{x+y}{2}, y]$.

3. Proof of the open induction principle

In this section we give the complete proof of the open induction principle for *enumerative* open sets.

Theorem 3.1 (Open Induction, enumerative version)**.** *Let A be an enumerative open set of the closed interval $[0,1]$, satisfying the following property (A is said to be inductive) :*

$$\forall x \in [0,1],\ [\forall y \in [0,x[, y \in A] \Rightarrow x \in A$$

Then $A = [0,1]$.

3.1. **Encoding.** Let G denote the family of open intervals defining the enumerative open set A :

$$G := \{]\alpha_i, \beta_i[\ ,\ i \in \mathbb{N}\} \text{ and } A = \bigcup_{i\in\mathbb{N}}]\alpha_i, \beta_i[$$

We define in a mutually recursive way a predicate *acceptable* over $\mathbb{N}^*$ and a function f which maps the elements of $\mathbb{N}^*$ to the closed subintervals of $[0,1]$ with dyadic bounds.

- ▶ Definition of *acceptable*:
  - ε is acceptable
  - $u \bullet 0$ is acceptable iff u is acceptable
  - $u \bullet (n+1)$ is acceptable iff u is acceptable and $f(u)_l$ is included in the finite union of intervals $\bigcup_{i<n+1}]\alpha_i, \beta_i[$.
- ▶ Definition of f:
  - $f(\varepsilon) := [0,1]$

– $f(u \bullet 0) := f(u)_l$

– $f(u \bullet (n+1)) := \begin{cases} f(u)_r \text{ if } u \bullet (n+1) \text{ is acceptable} \\ f(u)_l \text{ otherwise} \end{cases}$

Remark 3.2. Let u be an element of $\mathbb{N}^*$ and $f(u) = [a, b]$.

- The width of $[a, b]$ is $\frac{1}{2^{|u|}}$.
- If u is a sequence of 0s, then u is acceptable.
- The closed interval $[0, a]$, possibly reduced to $\{0\}$, is included in A, as it is included in a finite union of open intervals from G.

We give in Figure 1, an example of a possible configuration of f and *acceptable*. Suppose that $[0, \frac{1}{2}]$ is included in $]\alpha_0, \beta_0[\bigcup]\alpha_1, \beta_1[\bigcup]\alpha_2, \beta_2[$ but not in $]\alpha_0, \beta_0[\bigcup]\alpha_1, \beta_1[$. On Figure 1, we have labeled certain nodes of T with the values of f and *acceptable* we can compute from that knowledge. A finite word is represented by the node it is the position of in the universal tree T. The first word (after 0) of length 1 to be acceptable will be the word 3, and every word n of length 1 with n greater than 3 will be *acceptable* as the union of elements of G goes on growing. Then 30 (the sequence consisting of 3 and 0) is *acceptable* as 3 itself is *acceptable*, and we have to check whether $[\frac{1}{2}, \frac{3}{4}]$ is included in $]\alpha_0, \beta_0[$ to compute $f(31)$.

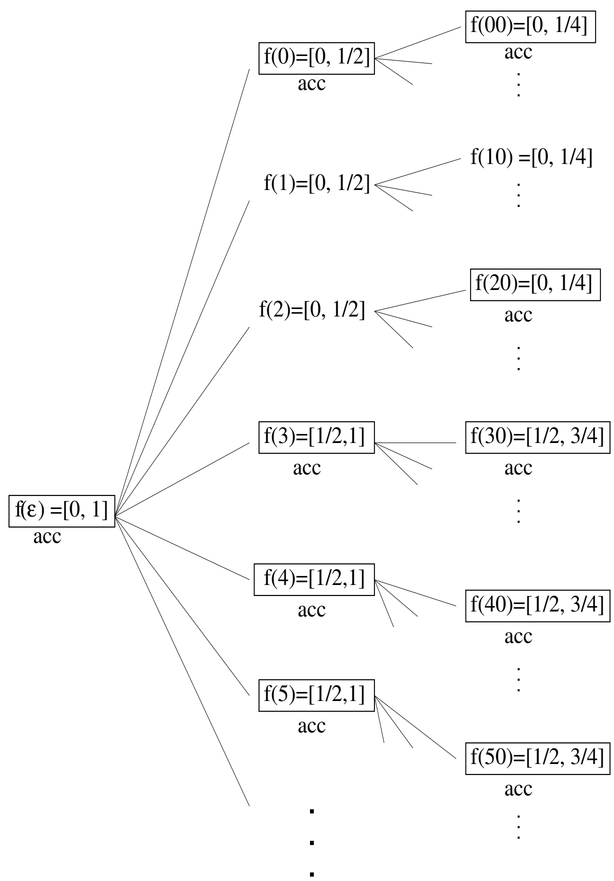

Figure 1. Example of configuration

3.2. **Bar lemma.** Let X be the predicate over $\mathbb{N}^*$ defined as:

$$X(u) \stackrel{\text{def}}{\equiv} u \text{ acceptable} \Rightarrow f(u) \text{ is included in } \bigcup_{i<|u|}]\alpha_i, \beta_i[.$$

The following lemma establishes that X is a bar.

Lemma 3.3. *For every infinite sequence $n_1 n_2 \dots$ in $\mathbb{N}^{\mathbb{N}}$, there exists k such that $f(n_1 \dots n_k)$ is included in $\bigcup_{i\leq k}]\alpha_i, \beta_i[$.*

Proof. We prove the following -stronger- proposition:
For every infinite sequence $n_1 n_2 \dots$ in $\mathbb{N}^{\mathbb{N}}$, there exists k such that $f(n_1 \dots n_k)$ is included in one of the $]\alpha_i, \beta_i[$ for $i = 0 \dots k$.

For all $k \in \mathbb{N}$, let $[a_k, b_k]$ denote $f(n_1 \dots n_k)$.The sequence $([a_k, b_k])_{k\in\mathbb{N}}$ is a sequence of shrinking nested closed intervals. Using the completeness of $\mathbb{R}$, one can assert that $\bigcap_{k\in\mathbb{N}} [a_k, b_k]$ is a singleton $\{x\}$. By Remark 3.2, for all $k \in \mathbb{N}$, $[0, a_k]$ is included in A, and so is $[0, y]$ for all $y \in [0, x[$. The set A being inductive, x is in A as well.

Consequently let n_0 be a natural number such that x belongs to $]\alpha_{n_0}, \beta_{n_0}[$. There exists k_0 such that $[a_{k_0}, b_{k_0}]$ is included in $]\alpha_{n_0}, \beta_{n_0}[$. Let k be the maximum of k_0 and n_0. As $k \geq k_0$, we have $[a_k, b_k] \subseteq [a_{k_0}, b_{k_0}]$, hence k is suitable.

□

3.3. **Using the bar induction axiom.** Let Y be the predicate over $\mathbb{N}^*$ defined as:

$$Y(u) \stackrel{\text{def}}{\equiv} u \text{ acceptable} \Rightarrow f(u) \text{ is included in a finite union of intervals of } G$$

For any finite word u, $X(u) \Rightarrow Y(u)$. Since all the prerequisite are gathered, we now prove using bar induction that the whole closed interval $[0, 1]$ is included in a finite union of intervals from G.

- For any finite word u, $X(u)$ implies $Y(u)$.
- X is monotonous: for every $n \in \mathbb{N}$, for every finite word u, $f(u \bullet n)$ is included in $f(u)$.
- Y is progressive: let u be a finite acceptable word. Let us assume moreover that for every $n \in \mathbb{N}$, $Y(u \bullet n)$. In particular, $f(u \bullet 0) = f(u)_l$ is included in a finite union of intervals from G. Let n be a natural number such that $f(u)_l$ is included in the finite union $\bigcup_{i<n+1}]\alpha_i, \beta_i[$. By definition, $u \bullet (n+1)$ is then acceptable and $f(u \bullet (n + 1)) = f(u)_r$ is included in a finite union of intervals from G. Thus, as both $f(u)_r$ and $f(u)_l$ are included in such a finite union, this is the case for $f(u)$.
- Lemma 3.3 asserts that X is a bar.

The bar induction principle proves $Y(\varepsilon)$ which is:
ε is acceptable $\Rightarrow$ $f(\varepsilon)$ is included in a finite union of open intervals from G
Yet as by definition ε is acceptable and $f(\varepsilon)$ is $[0, 1]$, we have proved that the *inductive* enumerative open set A contains the closed interval $[0, 1]$.

4. Computational content

The computational content of the proof uses *divide and conquer* like strategy. We start with a closed interval, $[0, 1]$, and an open set A which is known through a recursive function g. The current closed interval (that we want to prove to be included in A) is $[0, 1]$ and we haven't computed any value of g yet. At each step of the computation, if we have not succeeded in proving the inclusion, we split the current interval into two equal parts, focus on the left half, *and* compute at the same time a new value of the defining function g.

Thus to prove that $[0, 1]$ is included in A,we first focus on the left half of the closed interval $[0, \frac{1}{2}]$, compute a first value $g(0)$, which gives $]\alpha_0, \beta_0[$, and try to include $[0, \frac{1}{2}]$ in $]\alpha_0, \beta_0[$. Either we succeed and then we can skip to the next interval $[\frac{1}{2}, 1]$ which becomes the new current interval, or we don't and then we repeat the splitting step : the current interval becomes $[0, \frac{1}{4}]$, we compute the next value $g(1)$ and try to include $[0, \frac{1}{4}]$ into $]\alpha_0, \beta_0[\bigcup]\alpha_1, \beta_1[$.

The predicate *acceptable* marks the relevant nodes of the tree T, that is to say the nodes coding for a relevant computation. When a node u is *acceptable*, the two valid computations you can perform are, splitting the interval and computing a new value for g (that is why $u \bullet 0$ will be *acceptable*), or skipping to the next current interval if you have succeeded in the current task (that is the condition under which $u \bullet (n + 1)$ is *acceptable* as well).

The algorithmic content of the proof is then only the computation of $\bigcup_{k \leq n}]\alpha_k, \beta_k[$ successively from $n = 0$ up to the point it covers $[0, 1]$ and bar induction ensures the termination of that process.

Acknowledgments. The author wishes to thank gratefully G.Dowek for many advises on this work and the present article and also Th.Coquand and B.Spitters for their helpful comments about the open induction principle.

References

[1] Th. Coquand. A note on the open induction principle. 1997.

[2] Herman Geuvers and Milad Niqui. Constructive reals in coq: Axioms and categoricity. In *Types for Proofs and Programs, International Workshop, TYPES 2000, Durham, UK, December 8-12, 2000, Selected Papers*, volume 2277 of *Lecture Notes in Computer Science*, pages 79–95. Springer Verlag, 2000.

[3] Jean-Claude Raoult. Induction of open properties. Research Report RR-0813, INRIA, 1988.

[4] A. S. Troelstra and D. van Dalen. *Constructivism in mathematics. Vol. I*, volume 121 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, 1988. An introduction.

[5] A. S. Troelstra and D. van Dalen. *Constructivism in mathematics. Vol. II*, volume 123 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, 1988. An introduction.

[6] A.S. Troelstra, editor. *Metamathematical Investigation of Intutitionnistic Arithmetic and Analysis*, volume 344 of *Lecture Notes in Mathematics*. Springer Verlag, 1973.

[7] Wim Veldman. Almost the fan theorem. *Nieuw Arch. Wiskd. (5)*, 2(4), 2001.