Necessarily non-analytic induction proofs

FLoV Logic Seminar

Anders Lundstedt[*] Eric Johannesson[†]

Department of Philosophy, Stockholm University

Gothenburg, May 2019

[*] anders.lundstedt@philosophy.su.se, anderslundstedt.com

[†] eric.johannesson@philosophy.su.se

A non-analytic induction proof

Define $f : \mathbb{N} \to \mathbb{N}$ such that $f(n)$ is the sum of the first n odd natural numbers:

$$\begin{aligned} f(0) &\coloneqq 0, \\ f(n+1) &\coloneqq f(n) + 2n + 1. \end{aligned}$$

That is, we have

$$\begin{aligned} f(0) &= 0, \\ f(1) &= 1, \\ f(2) &= 1 + 3, \\ f(3) &= 1 + 3 + 5, \\ f(4) &= 1 + 3 + 5 + 7, \\ &\vdots \end{aligned}$$

A non-analytic induction proof

Fact

$f(n)$ is a perfect square for all n: For all natural numbers n there is a natural number m such that $f(n) = m^2$.

Let us try to prove this by “straightforward induction”; that is, let us try to prove the following.

- Base case: $f(0)$ is a perfect square.
- Induction step: For all natural numbers n, if $f(n)$ is a perfect square then $f(n+1)$ is a perfect square.

A non-analytic induction proof

Proof attempt of the induction step.

- Let n be any natural number.
- Induction hypothesis: There is a natural number k such that $f(n) = k^2$.
- We want to prove that $f(n+1) = m^2$ for some natural number m.
- We have

$$\begin{aligned} f(n+1) &= f(n) + 2n + 1 && \text{(by definition)} \\ &= k^2 + 2n + 1 && \text{(by induction hypothesis)} \end{aligned}$$

but $k^2 + 2n + 1$ is not a perfect square for arbitrary natural numbers k and n so how do we proceed from here?

A non-analytic induction proof

Let us try a different approach. Our fact follows immediately from the following stronger fact.

Fact

$f(n) = n^2$ for all natural numbers n.

(This fact is stronger in the sense that it logically implies the previous fact, while the previous fact does not logically imply this fact.)

Let us try to prove this fact by “straightforward induction”; that is, let us try to prove the following.

- Base case: $f(0) = 0^2$.
- Induction step: For all natural numbers n, if $f(n) = n^2$ then $f(n+1) = (n+1)^2$.

A non-analytic induction proof

Proof of the induction step.

- Let n be any natural number.
- Induction hypothesis: $f(n) = n^2$.
- We want to prove that $f(n+1) = (n+1)^2$.
- We have

$$\begin{aligned} f(n+1) &= f(n) + 2n + 1 && \text{(by definition)} \\ &= n^2 + 2n + 1 && \text{(by induction hypothesis)} \\ &= (n+1)^2. \end{aligned}$$

□

Terminology

- Proofs like these are commonly called something like "proof by a strengthening of the induction hypothesis".
- The typical form of a "straightforward induction proof":

$$\dfrac{\begin{array}{c}\vdots\\ \varphi(0)\end{array}\qquad\begin{array}{c}\vdots\\ \forall x\colon\ \varphi(x)\to\varphi(x+1)\end{array}}{\forall x.\,\varphi(x)}\quad .$$

- The typical form of a "proof by a strengthening of the induction hypothesis":

$$\begin{array}{c}\dfrac{\begin{array}{c}\vdots\\ \psi(0)\end{array}\qquad\begin{array}{c}\vdots\\ \forall x\colon\ \psi(x)\to\psi(x+1)\end{array}}{\forall x.\,\psi(x)}\\ \vdots\\ \forall x.\,\varphi(x)\end{array}\quad .$$

Terminology

There need not always be any precise sense in which $\psi(x)$ is stronger than $\varphi(x)$. Thus, following Hetzl and Wong, we use the more general terminology “non-analytic induction proofs”.[1]

[1] Stefan Hetzl and Tin Lok Wong (2018): “Some observations on the logical foundations of inductive theorem proving”.

The problem

- Question: Take a non-analytic induction proof (for example, the proof we just saw). Is the non-analyticity of this proof necessary?
- It is not immediately obvious how to make precise sense of this question. For example, if we would use the previously given forms to distinguish analytic induction proofs from non-analytic induction proofs, then any proof of $\forall x.\, \varphi(x)$ could be turned into an analytic induction proof:

$$\dfrac{\dfrac{\begin{matrix}\vdots\\ \forall x.\,\varphi(x)\end{matrix}}{\varphi(0)} \qquad \begin{matrix}\vdots\\ \forall x.\,\varphi(x)\\ \vdots\\ \forall x\colon\ \varphi(x)\to\varphi(x+1)\end{matrix}}{\forall x.\,\varphi(x)}\ .$$

The problem

- Hetzl and Wong have made precise nontrivial sense of the notion of “necessarily non-analytic induction proof”.
- Our main result so far: Using a slight generalization of Hetzl’s and Wong’s formulation, there is a precise sense in which we must use non-analytic induction to prove “the sum of any initial segment of the odd natural numbers is a perfect square”.

Motivation

Why try to make precise sense of the notion of "necessarily non-analytic induction proofs"? Why try to settle whether some non-analytic induction proofs are necessarily non-analytic?

- Curiosity!
- Make precise sense of imprecise things said or thought in mathematical practice.
- Show that we can make precise nontrivial sense of a subclass of the class of questions of the form "must one prove a fact from the class C of facts in order to prove the fact X?"
- Show that current "logical grounding" frameworks do not model mathematical justification and explanation.

Definitions

Definition

- The *full (first-order) language of arithmetic*, notation $\mathcal{L}_{\text{full}}$, is the first-order language that for each natural number n has
    - a constant symbol n,
    - a function symbol f of arity $n+1$ for each function $f : \mathbb{N}^{n+1} \to \mathbb{N}$,
    - a relation symbol P of arity n for each relation $P \subseteq \mathbb{N}^n$.
- The *minimal (first-order) language of arithmetic*, notation $\mathcal{L}_{\text{min}}$, is the $\mathcal{L}_{\text{full}}$-reduct with signature $\langle 0, 1, + \rangle$.
- A first-order language L is a *(first-order) language of arithmetic* if and only if L is an $\mathcal{L}_{\text{min}}$-expansion and an $\mathcal{L}_{\text{full}}$-reduct.

Definitions

Definition

Let L be a language of arithmetic.

- The *standard L-model* has domain $\mathbb{N}$ and each symbol interpreted as itself.
- An L-model not isomorphic to the standard L-model is *non-standard*.
- The L-theory *true L-arithmetic* is the theory of the standard L-model.
- Any subset of true L-arithmetic is a *theory of L-arithmetic*.

Definitions

Definition

Let L be a language of arithmetic and let $\varphi(x)$ be an L-formula in the free variable x. The *induction instance* for $\varphi(x)$ is the L-sentence

$$\mathrm{IND}(\varphi) :\equiv \varphi(0) \wedge \forall x(\varphi(x) \to \varphi(x+1)) \to \forall x.\, \varphi(x).$$

Definitions

Definition

Let L be a language of arithmetic. Let T be an L-theory. Let $\varphi(x)$ be an L-formula in the free variable x. *T proves $\forall x.\, \varphi(x)$ by necessarily non-analytic induction* if and only if there is an L-formula $\psi(x)$ in the free variable x such that

$$\begin{aligned}
&(1) & T, \mathrm{IND}(\varphi) &\nvdash \forall x.\, \varphi(x),\\
&(2) & T &\vdash \varphi(0),\\
&(3) & T &\vdash \psi(0),\\
&(4) & T &\vdash \forall x\colon\ \psi(x) \rightarrow \psi(x+1),\\
&(5) & T &\vdash \forall x.\, \psi(x) \rightarrow \forall x.\, \varphi(x).
\end{aligned}$$

Under conditions (1)–(5), we also say that $\psi(x)$ *witnesses* that T proves $\forall x.\, \varphi(x)$ by necessarily non-analytic induction.

Definitions

Let $\mathcal{L}^{\mathrm{OR}}$ be the language of ordered rings—signature $\langle 0, 1, +, \cdot, < \rangle$. We find it very reasonable that working mathematicians take the axioms of the $\mathcal{L}^{\mathrm{OR}}$-theory PA^-—*the theory of the non-negative parts of nontrivial discretely ordered commutative rings*[2]—for granted when doing arithmetic.

The axioms of PA^- are:

- associativity of addition: $(x + y) + z = x + (y + z)$,
- associativity of multiplication: $(x \cdot y) \cdot z = x \cdot (y \cdot z)$,
- commutativity of addition: $x + y = y + x$,
- commutativity of multiplication: $x \cdot y = y \cdot x$,
- distributivity of multiplication over addition: $x \cdot (y + z) = x \cdot y + x \cdot z$,

[2] As introduced in for example Richard Kaye's *Models of Peano Arithmetic* (1991).

Definitions

The axioms of PA$^-$, continued:

- 0 is an additive identity: $x + 0 = 0$,
- 0 is a multiplicative zero: $x \cdot 0 = 0$,
- 1 is a multiplicative identity: $x \cdot 1 = x$,
- the order is irreflexive: $x \not< x$,
- the order is transitive: $x < y \wedge y < z \to x < z$,
- the order is total: $x < y \vee x = y \vee y < x$,
- addition respects the order: $x < y \to x + z < y + z$,
- multiplication respects the order: $0 < z \wedge x < y \to x \cdot z < y \cdot z$,
- smaller elements can be subtracted from larger elements: $x < y \to \exists z.\, x + z = y$,
- $0 < 1$,
- the order is discrete: $0 < x \to x = 1 \vee 1 < x$,
- 0 is the least element: $x = 0 \vee 0 < x$.

Our result

- Recall that we defined $f : \mathbb{N} \to \mathbb{N}$ such that $f(n)$ is the sum of the n first odd natural numbers.
- Recall that a language of arithmetic (as we defined it) has as symbols natural numbers and functions and relations on natural numbers.
- Expand $\mathcal{L}^{\mathrm{OR}}$ to a language L of arithmetic by adding f as a function symbol.

Our result

- Expand PA^- to a theory T of L-arithmetic by adding the defining equations for f:

$$T \coloneqq \mathrm{PA}^- \cup \{f(0) = 0,\ \forall x.\, f(x+1) = f(x) + 2x + 1\}.$$

- Define L-formulas $\varphi(x)$ and $\psi(x)$ corresponding to the analytic and non-analytic induction hypotheses, respectively:

$$\varphi(x) :\equiv \exists y.\, f(x) = y^2,$$
$$\psi(x) :\equiv f(x) = x^2.$$

Fact

$\psi(x)$ witnesses that T proves $\forall x.\, \varphi(x)$ by necessarily non-analytic induction.

Proof of our result

- Conditions (2)–(5) are easy to show.
- To show condition (1),

$$T, \mathrm{IND}(\varphi) \nvdash \forall x.\, \varphi(x),$$

we exhibit a non-standard L-model $M \models T$ with a non-standard number c such that

$$\begin{aligned} M &\models \varphi(c), \\ M &\not\models \varphi(c+1). \end{aligned}$$

Proof of our result

- $\mathbb{Z}[X] := \langle \mathbb{Z}[X], 0, 1, +, \cdot, < \rangle$ is the ordered ring of polynomials in the indeterminate X with coefficients in $\mathbb{Z}$.
- Elements of $\mathbb{Z}[X]$ are polynomials

$$z_n X^n + \cdots + z_1 X^1 + z_0$$

with $z_0, \ldots, z_n$ in $\mathbb{Z}$ and if $n \neq 0$ then $z_n \neq 0$. n is called the *degree* of the polynomial.

Proof of our result

- Addition, multiplication and subtraction in $\mathbb{Z}[X]$ are as expected.
- The order can be thought of as given by taking X to be infinitely large, and taking X^{n+1} to be infinitely larger than X^n for each natural number n. Making this precise, we may define the order by the clauses

$$\begin{aligned} z_nX^n + \cdots + z_1X^1 + z_0 > 0 &\text{ if and only if } z_n > 0, \\ p > q &\text{ if and only if } p - q > 0. \end{aligned}$$

Proof of our result

- The polynomials in $\mathbb{Z}[X]$ can be divided into the *constant* polynomials

  $$z \quad (z \text{ in } \mathbb{Z})$$

  and the *non-constant* polynomials

  $$pX + z \quad (p \text{ in } \mathbb{Z}[X],\ p \neq 0,\ z \text{ in } \mathbb{Z}).$$

Proof of our result

- Let $\mathbb{Z}[X]^+$ be the non-negative part of $\mathbb{Z}[X]$; that is, $\mathbb{Z}[X]^+$ is the substructure of $\mathbb{Z}[X]$ that consists of polynomials of the form

$$z_n X^n + \cdots + z_1 X^1 + z_0$$

with $z_n \geq 0$ (and $z_n = 0$ only if $n = 0$).

Proof of our result

Fact

An $\mathcal{L}^{OR}$-model M is a model of PA^- if and only if M is the non-negative part of a nontrivial discretely ordered commutative ring.

Proof.

See for example Kaye's *Models of Peano Arithmetic.* □

Corollary

$\mathbb{Z}[X]^+ \models PA^-$.

Proof.

$\mathbb{Z}[X]^+$ is the non-negative part of the nontrivial discretely ordered commutative ring $\mathbb{Z}[X]$. □

Proof of our result

- We want to expand $\mathbb{Z}[X]^+$ to an L-model $M \vDash T$ such that $M \vDash \varphi(p)$ and $M \nvDash \varphi(p+1)$ for some polynomial p in $\mathbb{Z}[X]^+$.
- To expand $\mathbb{Z}[X]^+$ to an L-model M we need to provide an interpretation $f^M : \mathbb{Z}[X]^+ \to \mathbb{Z}[X]^+$ of f.
- Recall that

  $$T = \mathrm{PA}^- \cup \{f(0) = 0,\ \forall x.\, f(x+1) = f(x) + 2x + 1\}.$$

  and that

  $$\varphi(x) :\equiv \exists y.\, f(x) = y^2.$$

- Thus f^M needs to satisfy the defining equations for f and be such that for some polynomial p in $\mathbb{Z}[X]^+$ we have that $f^M(p)$ is a perfect square in $\mathbb{Z}[X]^+$ while $f^M(p+1)$ is not.

Proof of our result

- Since M must model the recursive defining equation,

$$f(x+1) = f(x) + 2x + 1,$$

we get that f^M must satisfy

$$\begin{aligned} f^M(p) &= f^M((p-1)+1) \\ &= f^M(p-1) + 2(p-1) + 1. \\ &= f^M(p-1) + 2p - 1. \end{aligned}$$

Thus f^M must satisfy

$$f^M(p-1) = f^M(p) - 2p + 1.$$

Proof of our result

- f^M must thus satisfy the equations
$$\begin{aligned} f^M(0) &= 0, \\ f^M(p+1) &= f^M(p) + 2p + 1, \\ f^M(p-1) &= f^M(p) - 2p + 1. \end{aligned}$$
- The first two equations fixes f^M on the constant polynomials.
- Let $pX + z$ be a non-constant polynomial in $\mathbb{Z}[X]^+$ and let q be any polynomial in $\mathbb{Z}[X]^+$. By the last two equations, setting $f^M(pX + z) = q$ fixes f^M on all polynomials of the form $pX + z'$ (z' in $\mathbb{Z}$), that is it fixes f^M on the polynomials
$$pX + z + 1,\ pX + z + 2,\ \ldots$$
$$pX + z - 1,\ pX + z - 2,\ \ldots$$

Proof of our result

- Thus what we need to do is: For each $p > 0$ in $\mathbb{Z}[X]^+$, define $f^M(pX + z)$ for some z in $\mathbb{Z}$;
- In doing so, making sure that
  - $f^M(p)$ is in $\mathbb{Z}[X]^+$ for all (non-constant) p in $\mathbb{Z}[X]^+$,
  - for some (non-constant) p in $\mathbb{Z}[X]^+$, $f^M(p)$ is a perfect square while $f^M(p + 1)$ is not.
- For each $p > 0$ in $Z[X]^+$, define

$$f^M(pX - 1) := pX^2.$$

Proof of our result

- $f^M(p)$ is in $\mathbb{Z}[X]^+$ for all (non-constant) p in $\mathbb{Z}[X]^+$: We need to worry about the equation
$$f^M(p-1) = f^M(p) - 2p + 1.$$
By construction, $f^M(p)$ is always positive and of greater degree than p for non-constant polynomials p. Thus the right hand side will never be negative.
- We have
$$f^M(X-1) = X^2$$
and
$$\begin{aligned} f^M(X) &= f(X-1) + 2(X-1) + 1 \\ &= X^2 + 2X - 1. \end{aligned}$$
Thus $f^M(X-1)$ is a perfect square in $\mathbb{Z}[X]^+$ while $f^M(X)$ is not. This completes the proof.

Summary of our result

Let

$$T := \mathrm{PA}^{-} \cup \{f(0) = 0,\ \forall x.\, f(x+1) = f(x) + 2x + 1\}.$$

and let

$$\begin{aligned} \varphi(x) &:\equiv \exists y.\, f(x) = y^2, \\ \psi(x) &:\equiv f(x) = x^2. \end{aligned}$$

Fact

$\psi(x)$ witnesses that T proves $\forall x.\, \varphi(x)$ by necessarily non-analytic induction.

Summary of our result

Proof.

- Conditions (2)–(5) are easy.
- To show condition (1),

  $$T, \mathrm{IND}(\varphi) \nvdash \forall x.\, \varphi(x),$$

  we exhibit a non-standard L-model $M \models T$ with a non-standard natural number c such that $M \models \varphi(c)$ and $M \not\models \varphi(c+1)$.
- $\mathbb{Z}[X]^+$ is a model of PA^-. We expand $\mathbb{Z}[X]^+$ to an L-model M by interpreting f on $\mathbb{Z}[X]^+$.
- We define our interpretation f^M such that it satisfies the defining equations for f and such that $f^M(X-1)$ is a perfect square in $\mathbb{Z}[X]^+$ while $f^M(X)$ is not. □

Ideas for future work

Our proof breaks down if we add any sentence to T that is false in $\mathbb{Z}[X]^+$. A natural such sentence that is true in the standard model is “all numbers are even or odd”, that is

$$\sigma :\equiv \forall x \exists y.\, x = y + y \vee x = y + y + 1.$$

Conjecture

$\psi(x)$ witnesses that $T \cup \{\sigma\}$ proves $\forall x.\, \varphi(x)$ by necessarily non-analytic induction.

Ideas for future work

- Develop more general methods to settle conjectures about necessary non-analyticity (as opposed to the method of hand-crafting countermodels for each particular case).
- Consider other settings than arithmetic. For example, in computer science, many basic facts of functions on inductive structures seem to require non-analytic induction proofs.
- Consider the problem of non-analytic induction proofs from the more proof-theoretical side. Dag Prawitz's recent "The concepts of proof and ground" might be useful.[3]
- Relate necessarily non-analytic induction proofs to "logical grounding". The same work by Dag Prawitz should be useful here as well.

[3] Dag Prawitz (2018): "The concepts of proof and ground", preprint.

Thanks!

Thanks for listening!