WHEN MUST ONE USE A NON-ANALYTIC INDUCTION HYPOTHESIS?——SUMMARY OF SOME RESULTS

version: 2019-03-15

Eric Johannesson and Anders Lundstedt

eric.johannesson@philosophy.su.se
anders.lundstedt@philosophy.su.se

§1 Introduction

In §2 we give two examples of facts that when proved seem to require "a strengthening of the induction hypothesis". The two facts are:

(F1) For all natural numbers n: $1+3+5+\cdots+(2n-1) = k^2$ for some natural number k.

(F2) For all natural numbers n: $1+1/2^2+1/3^2+\cdots+1/(n+1)^2 < 2$.

**

Hetzl and Wong (2017) have made precise sense of the notion of "proof by using a non-analytic induction hypothesis". (What one might call "proof by a strengthening of the induction hypothesis" is a special case of "proof by using a non-analytic induction hypothesis".) In §3 we present a slight generalization of their formalization.

**

In §4 we show that, using the definition given in §3, there is a precise sense in which (F1) must be proved by "using a non-analytic induction hypothesis".

**

In §5 we present some preliminary results towards proving (or disproving) that (F2), in the sense given by the definition in §3, must be proved by "using a non-analytic induction hypothesis".

**

In §6 we present some ideas for future work.

**

§2 Two results proved by "a strengthening of the induction hypothesis"

For our first case, we have that the sum of any initial segment of the odd numbers is a perfect square, that is,

> for all natural numbers n: $1+3+5+\cdots+(2n-1) = k^2$ for some natural number k,

or, avoiding ellipsis notation, the following.

**

DEFINITION. $f_1 : \mathbb{N} \to \mathbb{N}$ is recursively defined by

$$f_1(0) := 0,$$

$$f_1(n+1) := f_1(n)+2n+1.$$

**

FACT 1. For all natural numbers n: $f_1(n) = k^2$ for some natural number k.

PROOF ATTEMPT. "Straightforward induction":

- Base case: $f_1(0) = 0 = 0^2$ (by definitions) so $f_1(0) = k^2$ for $k = 0$.

- Induction step:

$$\begin{aligned} f_1(n+1) &= f_1(n)+2n+1 && \text{(by definition)} \\ &= k^2+2n+1 && \text{(for some k, by induction hypothesis).} \end{aligned}$$

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

ACTUAL PROOF. It suffices to prove the following "stronger" fact, of which Fact 1 is a logical consequence.

> $f_1(n) = n^2$ for all natural numbers n.

(This fact is not a logical consequence of Fact 1, so it is stronger than Fact 1 in at least that sense.)

We prove this fact by induction.

- Base case: $f_1(0) = 0^2 \equiv 0 = 0$ (by definitions).

- Induction step:

$$\begin{aligned} f_1(n+1) &= f_1(n)+2n+1 && \text{(by definition)} \\ &= n^2+2n+1 && \text{(by induction hypothesis)} \\ &= (n+1)^2 && \text{(by elementary arithmetic).} \end{aligned}$$ □

**

For our second case, we have

> $1+1/2^2+1/3^2+\cdots+1/(n+1)^2 < 2$ for all natural numbers n,

or, avoiding ellipsis notation, the following.

**

DEFINITION. $f_2 : \mathbb{N} \to \mathbb{Q}$ is recursively defined by

$$\begin{aligned} f_2(0) &:= 1, \\ f_2(n+1) &:= f_2(n)+1/(n+2)^2. \end{aligned}$$

**

FACT 2. $f_2(n) < 2$ for all natural numbers n.

PROOF ATTEMPT. "Straightforward" induction:

- Base case: $f_2(0) < 2 \equiv 1 < 2$ (by definition).

- Induction step:

$$\begin{aligned} f_2(n+1) &= f_2(n)+1/(n+2)^2 && \text{(by definition)} \\ &< 2+1/(n+2)^2 && \text{(by induction hypothesis).} \end{aligned}$$

  But $2+1/(n+2)^2 \not< 2$ for any natural number n so how do we proceed from here?

ACTUAL PROOF. It clearly suffices to prove the "stronger" fact

$$f_2(n) \leq 2-1/(n+1) \text{ for all natural numbers n.}$$

(In what sense is this fact stronger than Fact 2? It is at least stronger in the sense that for arbitrary $f : \mathbb{N} \to \mathbb{Q}$, $f(n) \leq 2-1/(n+1)$ for all n implies $f(n) < 2$ for all n while the converse implication need not hold.)

We prove this fact by induction.

- Base case: $f_2(0) \leq 2-1/(0+1) \equiv 1 \leq 1$ (by definitions).

- Induction step:

$$\begin{aligned} f_2(n+1) &= f_2(n)+1/(n+2)^2 && \text{(by definition)} \\ &\leq 2-1/(n+1)+1/(n+2)^2 && \text{(by induction hypothesis)} \\ &= 2-1/(n+2)-1/(n+1)(n+2)^2 && \text{(by lots of elementary arithmetic)} \\ &\leq 2-1/(n+2) && \text{(by more or less obvious facts).} \ \square \end{aligned}$$

**

§3 Definitions

What follows in this section is a reformulation and slight generalization of some of the notions introduced by Hetzl and Wong (2017).

**

DEFINITION. The *full (first-order) language of arithmetic*, notation $\mathcal{L}[\mathrm{full}]$, 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 $P \subseteq \mathbb{N}^n$.

**

DEFINITION. The *minimal (first-order) language of arithmetic*, notation $\mathcal{L}[\mathrm{min}]$, is the $\mathcal{L}[\mathrm{full}]$-reduct with signature $\langle 0,1,+,\cdot,< \rangle$.

**

DEFINITION. A first-order language L is a *(first-order) language of arithmetic* if and only if L is an $\mathcal{L}[\mathrm{min}]$-expansion and an $\mathcal{L}[\mathrm{full}]$-reduct.

**

Thus a first-order language of arithmetic has as symbols natural numbers and operations on natural numbers and relations on natural numbers. This is just a convenient "trick" which allows us to easily define the standard model.

**

DEFINITION. Let L be a language of arithmetic.

- The *standard L-model* has domain $\mathbb{N}$ and each symbol interpreted as itself.

- 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*.

**

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

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

**

DEFINITION. Let L be a language of arithmetic and let T be an L-theory. Let $\varphi(x)$ and $\psi(x)$ be L-formulas both with at most one free variable x. Say that *ψ witnesses that T proves $\forall x.\varphi(x)$ with and only with a non-analytic induction hypothesis* if and only if

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

(2) $T \vdash \varphi(0)$,

(3) $T \vdash \psi(0)$,

(4) $T \vdash \forall x,\psi(x)\to\psi(x+1)$,

(5) $T \vdash \forall x.\psi(x)\to\forall x.\varphi(x)$.

**

DEFINITION. The £[min]-theory PA^- is axiomatized by the following.

- 0 and 1 are distinct:

  $$0 \neq 1.$$

- Associativity of addition and multiplication:

  $$(x+y)+z = x+(y+z),$$

  $$(x\cdot y)\cdot z = x\cdot (y\cdot z).$$

- Commutativity of addition and multiplication:

  $$x+y = y+x,$$

  $$x\cdot y = y\cdot x.$$

- Distributivity of addition over multiplication:

  $$x\cdot (y+z) = x\cdot y+x\cdot z.$$

- 0 is an additive identity and a multiplicative zero:

  $$x\cdot 0 = 0.$$

- The order is irreflexive:

  $$x \not< x.$$

- The order is transitive:

  $$x < y \wedge y < z \rightarrow x < z.$$

- The order is total:

  $$x < y \vee x = y \vee y < x.$$

- The order is discrete:

  $$x = 0 \vee x = 1 \vee 1 < x.$$

- Addition and multiplication respect the order:

  $$x < y \rightarrow x+z < y+z,$$

  $$x < y \wedge 0 \neq z \rightarrow x\cdot z < y\cdot z.$$

- Smaller elements can be subtracted from larger:

  $$x < y \rightarrow \exists z.\ x+z = y.$$

**

FACT. For all £[min]-structures M: $M \models PA^-$ if and only if M is the non-negative part of a nontrivial discretely ordered commutative ring.

**

FACT. PA^- is a theory of arithmetic.

**

FACT. $PA = PA^- \cup \{IND(\varphi) : \varphi$ an $\mathcal{L}[min]$-formula with at most one free variable$\}$

**

§4 Fact 1 must be proved using a non-analytic induction hypothesis

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

**

Elements of $\mathbb{Z}[X]$ are polynomials

$$z_0+z_1X+z_2X^2+\cdots+z_nX^n$$

with $z_0,z_1,\ldots,z_n$ in $\mathbb{Z}$ and if $n \neq 0$ then $z_n \neq 0$. These can be divided into the *constants* polynomials

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

and the *non-constant* polynomials

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

**

Addition and multiplication in $\mathbb{Z}[X]$ are as expected. The order is given by

$$z_0+z_1X+z_2X^2+\cdots+z_nX^n > 0 \text{ if and only if } z_n > 0,$$

$$p > q \text{ if and only if } p-q > 0.$$

**

DEFINITION. $\mathbb{Z}[X]^+ = \langle \mathbb{Z}[X]^+,0,1,+,\cdot,< \rangle$ is the non-negative part of $\mathbb{Z}[X]$, that is, a polynomial p from $\mathbb{Z}[X]$ is in $\mathbb{Z}[X]^+$ if and only if $p \geq 0$.

**

Elements of $\mathbb{Z}[X]^+$ are polynomials

$$z_0+z_1X+z_2X^2+\cdots+z_nX^n$$

with $z_0,z_1,\ldots,z_n$ in $\mathbb{Z}$ and $z_n \geq 0$ and if $n \neq 0$ then $z_n \neq 0$. The constant polynomials are

$$n \quad (n \text{ in } \mathbb{N})$$

and the non-constant polynomials are

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

**

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

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

**

DEFINITION. The language of arithmetic L_1 is $£[min]$ expanded with the function symbol f_1.

**

DEFINITION. The theory of arithmetic T_1 is defined by

$$T_1 := PA^-,\ f_1(0) = 0,\ \forall x.\ f_1(x+1) = f_1(x)+2x+1.$$

**

DEFINITION. The L-formulas $\varphi_1(x)$ and $\psi_1(x)$ are defined by

$$\varphi_1(x) :\equiv \exists y.\ f_1(x) = y\cdot y,$$

$$\psi_1(x) :\equiv f_1(x) = x\cdot x.$$

**

FACT. ψ_1 witnesses that T_1 proves $\forall x.\varphi_1(x)$ with and only with a non-analytic induction hypothesis.

PROOF. We need to verify conditions (1)-(5).

(2) Trivial.

(3) Trivial.

(4) The proof of Fact 1 given earlier can be straightforwardly formalized in PA^-.

(5) Trivial.

(1) We exhibit an L_1-model satisfying T_1 and $IND(\varphi_1)$ but not $\forall x.\varphi_1(x)$. Let M be the L_1-expansion of $\mathbb{Z}[X]^+$ with f_1 interpreted as follows.

$$\begin{aligned} f_1{}^M(0) &:= 0,\\ f_1{}^M(n+1) &:= f_1{}^M(n)+2n+1,\\ f_1{}^M(pX-1) &:= pX^2\\ f_1{}^M(pX-1+(n+1)) &:= f_1{}^M(pX-1+n)+2(pX-1+n)+1,\\ f_1{}^M(pX-1-(n+1)) &:= f_1{}^M(pX-1-n)-2(pX-1-n)+1. \end{aligned}$$

The right hand side of the last equation does indeed define a polynomial of $\mathbb{Z}[X]^+$ since by construction the degree of $f_1^M(z+pX)$ is greater than the degree of pX for all integers z and all p in $\mathbb{Z}[X]^+$. (We get the last equation by solving

$$f_1(pX-1-n) = f_1(pX-1-(n+1)+1) = f_1(pX-1-(n+1))+2(pX-1-(n+1))+1$$

for $f_1(pX-1-(n+1))$.)

We then have

$$f_1^M(X-1) = X^2$$

which is a perfect square in M and we have

$$\begin{aligned} f_1^M(X) &= f_1^M(X-1)+2(X-1)+1 \\ &= X^2+2X-1 \end{aligned}$$

which is not a perfect square in M. Thus:

- $M \nvDash \forall x.\varphi_1(x)$ since $M \nvDash \varphi_1(X)$.
- $M \vDash \mathrm{IND}(\varphi_1)$ since $M \nvDash \forall x,\varphi_1(x)\to\varphi_1(x+1)$ since $M \vDash \varphi_1(X)$ but $M \nvDash \varphi_1(X+1)$.

By construction we also have $M \vDash T_1$ so we are done. □

**

§5 Must Fact 2 be proved using a non-analytic induction hypothesis?

Since Fact 2 is a statement involving rationals, a little work is needed to phrase it as a natural statement in first-order arithmetic (that is, without involving any coding).

**

DEFINITION. $g : \mathbb{N} \to \mathbb{N}$ and $h : \mathbb{N} \to \mathbb{N}$ are recursively defined by

$$g(0) := 1,$$

$$g(n+1) := (n+2)^2 g(n)+h(n),$$

$$h(0) := 1,$$

$$h(n+1) := (n+2)^2 h(n).$$

**

We have

$$f_2(n) = g(n)/h(n)$$

so the inequality $f_2(n) < 2$ can be rewritten as

$$g(n) < 2h(n).$$

Similarly the inequality $f_2(n) \leq 2-1/(n+1)$ becomes

$$(n+1)g(n) \leq (2n+1)h(n).$$

**

All in all, we have the following rephrasings of Fact 2 and its proof attempt and proof.

**

LEMMA. For all natural numbers n, if $(n+1)g(n) \leq (2n+1)h(n)$ then $g(n) < 2h(n)$.

PROOF. Suppose $(n+1)g(n) \leq (2n+1)h(n)$. Since $(2n+1)h(n) < 2(n+1)h(n)$ we then have $(n+1)g(n) < 2(n+1)h(n)$ so we must have $g(n) < 2h(n)$. □

**

FACT 2'. $g(n) < 2h(n)$ for all n.

PROOF ATTEMPT. Induction on n.

- Base case:

$$\begin{aligned} g(0) < 2h(0) &\equiv 1 < 2\cdot 1 \quad \text{(by definition)} \\ &\equiv 1 < 2. \end{aligned}$$

- Induction step:

$$\begin{aligned} g(n+1) &= (n+2)^2 g(n)+h(n) && \text{(by definition)} \\ &< (n+2)^2\cdot 2h(n)+h(n) && \text{(by induction hypothesis)} \end{aligned}$$

But for any n we have $(n+2)^2\cdot 2h(n)+h(n) \not< 2h(n)$. Thus we are "stuck".

ACTUAL PROOF. By the lemma, it suffices to prove

$$(n+1)g(n) \leq (2n+1)h(n) \text{ for all n.}$$

We prove this by induction on n.

- Base case:

$$\begin{aligned} (0+1)g(0) \leq (2\cdot 0+1)h(0) \quad &\text{(by definition)} \\ \equiv 1\cdot 1 \leq 1\cdot 1 & \\ \equiv 1 \leq 1. & \end{aligned}$$

- Induction step: Our induction hypothesis is

  (IH) $(n+1)g(n) \leq (2n+1)h(n)$

  and we must show

  (GOAL) $(n+2)g(n+1) \leq (2n+3)h(n+1)$.

  We have

$$\begin{aligned}(n+2)g(n+1) &= (n+2)((n+2)^2g(n)+h(n)) \\ &= (n+2)^3g(n)+(n+2)h(n)\end{aligned}$$

  and

$$\begin{aligned}(2n+3)h(n+1) &= (2n+3)(n+2)^2h(n) \\ &= (2n+3)(n+2)(n+2)h(n) \\ &= (2n^2+7n+6)(n+2)h(n) \\ &= (2n^2+7n+5)(n+2)h(n)+(n+2)h(n)\end{aligned}$$

  so (GOAL) is equivalent to

$$(n+2)^3g(n) \leq (2n^2+7n+5)(n+2)h(n),$$

  which is equivalent to

$$(n+2)^2g(n) \leq (2n^2+7n+5)h(n),$$

  which we have by

$$\begin{aligned}(n+2)^2g(n) &= ((n+1)+1)^2g(n) \\ &= ((n+1)^2+2(n+1)+1)g(n) \\ &= ((n+1)^2+2(n+1))g(n)+g(n) \\ &= (n+3)(n+1)g(n)+g(n) \\ &\leq (n+3)(2n+1)h(n)+g(n) && \text{(by (IH))} \\ &\leq (n+3)(2n+1)h(n)+2h(n) && \text{(by Lemma and (IH))} \\ &= (2n^2+7n+5)h(n).\end{aligned}$$ □

**

DEFINITION. The language of arithmetic L_2 is $\mathcal{L}$[min] expanded with the function symbols g and h.

**

DEFINITION. The L_2-sentences DEF(g) and DEF(h) and the L_2-formulas $\varphi_2(x)$ and $\psi_2(x)$ are defined by

$$\mathrm{DEF}(g) :\equiv g(0) = 1 \land \forall x.\ g(x+1) = (x+2)^2\cdot g(x)+h(x)$$

$$\mathrm{DEF}(h) :\equiv h(0) = 1 \land \forall x.\ h(x+1) = (x+2)^2\cdot h(x),$$

$$\varphi_2(x) :\equiv g(x) < 2\cdot h(x),$$

$$\psi_2(x) :\equiv (x+1)\cdot g(x)+h(x) \leq 2\cdot(x+1)\cdot h(x).$$

**

CONJECTURE. There is an L_2-theory of arithmetic $T \supseteq PA^-$ such that

(1) $T \vdash DEF(g)$,

(2) $T \vdash DEF(h)$,

(3) ψ_2 witnesses that T proves $\forall x.\varphi_2(x)$ with and only with a non-analytic induction hypothesis.

**

Adapting the proof of the previous section——cleverly interpreting the new function symbols on $\mathbb{Z}[X]^+$——will not work to settle the above conjecture, as the following results shows.

**

LEMMA. Let M be an L_2-expansion of $\mathbb{Z}[X]^+$ such that $M \models DEF(h)$. Let p be a non-zero polynomial and let z be an integer. If $h^M(pX+z) \neq 0$ then $h^M(pX+z-1) \neq 0$ and $\text{degree}(h^M(pX+z)) > \text{degree}(h^M(pX+z-1))$.

PROOF. We have

$$h^M(pX+z) = (pX+z+1)^2 h^M(pX+z-1)$$

since $M \models DEF(h)$ so clearly if $h^M(pX+z) \neq 0$ then $h^M(pX+z-1) \neq 0$ and $\text{degree}(h^M(pX+z)) > \text{degree}(h^M(pX+z-1))$. □

**

FACT. Let M be an L_2-expansion of $\mathbb{Z}[X]^+$. If $M \models DEF(h)$ then $h(p) = 0$ for all non-constant polynomials p.

PROOF. Suppose $h(pX+z) \neq 0$ for some non-constant polynomial $pX+z$. By the above lemma we then have the contradiction that we have an infinite descending chain

$$\begin{aligned}\text{degree}(h^M(pX+z)) &> \text{degree}(h^M(pX+z-1))\\ &> \text{degree}(h^M(pX+z-2))\\ &> \text{degree}(h^M(pX+z-3))\\ &\vdots\end{aligned}$$ □

**

Note that the above results should generalize to any polynomial ring $R[X]$. Perhaps one could settle the conjecture by considering some ring $R[X,X^{-1}]$ of Laurent polynomials.

**

§5 Future work

One line of future work would of course be to settle the conjecture in §4.

**

In §3 we expanded $\mathbb{Z}[X]^+$ to a L_1-model in order to show that T_1 proves $\forall x.\varphi_1(x)$ with and only with a non-analytic induction hypothesis. For each sentence σ of true L_1-arithmetic that is false in $\mathbb{Z}[X]^+$ it is natural to ask whether adding σ to T_1 lets us prove $\forall x.\varphi_1(x)$ without needing a non-analytic induction hypothesis. One simple sentence of true L_1-arithmetic that is not true in $\mathbb{Z}[X]^+$ is "all numbers are odd or even", that is

$$\sigma_1 :\equiv \forall x \exists y,\ x = y+y \lor x = y+y+1.$$

**

CONJECTURE. ψ_1 witnesses that $T_1 \cup \{\sigma_1\}$ proves $\forall x.\varphi_1(x)$ with and only with a non-analytic induction hypothesis.

**

To more systematically settling conjectures like the one above one could attempt to establish more general results——instead of hand-crafting countermodels for each particular case. We hope that the literature on first-order arithmetic already contains lots of applicable results.

**

Veryfing provability in weak fragments of arithmetic is hard. It is easy to rely on something true that is not provable in the fragment one works with. Thus it would be worthwhile to verify the provability statements in §3 with a theorem prover.

**

An interesting future line of work would be to consider other settings than arithmetic. For example, in computer science, basic facts of operations on inductive structures often seem to require a non-analytic induction hypothesis.

**

One might also approach the problem of non-analytic induction proofs from the more proof-theoretical side, for example by studying derivations in natural deduction. Dag Prawitz's (2018) recent work may be useful.

**

REFERENCES

Hetzl, Stefan and Tin Lok Wong (2017): "Some observations on the logical foundations of inductive theorem proving", Logical Methods in Computer Science 13(4).

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