Studying the role of induction axioms in reverse mathematics

Paul Shafer
Universiteit Gent
Paul.Shafer@UGent.be
http://cage.ugent.be/~pshafer/

SLS 2014
February 21, 2014

Contents

Let's begin with an old friend

The system RCA_0 is axiomatized by

- the axioms of a discretely ordered commutative semi-ring with 1 (or PA^- if you like);
- the Δ^0_1 comprehension scheme:

$$\forall n(\varphi(n) \leftrightarrow \psi(n)) \rightarrow \exists X \forall n(n \in X \leftrightarrow \varphi(n)),$$

  where φ is Σ^0_1 and ψ is Π^0_1; and
- the Σ^0_1 induction scheme:

$$(\varphi(0) \wedge \forall n(\varphi(n) \rightarrow \varphi(n+1))) \rightarrow \forall n \varphi(n),$$

  where φ is Σ^0_1.

Idea: To get new sets, you have to compute them from old sets.

Boy, lots of theorems say that certain kinds of sets exist!

RCA_0 gives you

- a little bit of comprehension and
- a little bit of induction.

In reverse mathematics, we typically analyze which theorems prove which other theorems over RCA_0.

Often a theorem looks like a set-existence principle: for every this kind of set, there is a that kind of set. For example:

- For every continuous function $[0, 1] \to \mathbb{R}$ there is a maximum.
- For every commutative ring there is a prime ideal.
- For every coloring of pairs in two colors there is a homogeneous set.
- For every infinite subtree of $2^{<\mathbb{N}}$ there is an infinite path.

A closer look at induction

So reverse mathematics is often concerned with set-existence principles.

Almost as often, induction merits no special consideration.
(But it can get tricky to make arguments work using only Σ^0_1-induction.)

Today, we look at a few situations in which induction plays a more pronounced role, such as with

- pigeonhole principles,
- conservativity results, and
- diagonally non-recursive functions.

It will be fun and interesting.

Induction and collection

The induction axiom for φ is (the universal closure of)

$$(\varphi(0) \wedge \forall n(\varphi(n) \to \varphi(n+1))) \to \forall n\varphi(n).$$

The collection (or bounding) axiom for φ is (the universal closure of)

$$(\forall n < t)(\exists m)\varphi(n,m) \to (\exists b)(\forall n < t)(\exists m < b)\varphi(n,m)$$

$\mathsf{I}\Sigma^0_n$ ($\mathsf{I}\Pi^0_n$) is the collection of induction axioms where the φ is Σ^0_n (Π^0_n).
Note RCA_0 includes $\mathsf{I}\Sigma^0_1$.

$\mathsf{B}\Sigma^0_n$ ($\mathsf{B}\Pi^0_n$) is the collection of bounding axioms where the φ is Σ^0_n (Π^0_n).

Induction and collection

Even though there is no such thing as a Δ^0_n formula, we can still make sense of $\mathsf{I}\Delta^0_n$.

$\mathsf{I}\Delta^0_n$ is the collection of universal closures of formulas of the form

$$\forall n(\varphi(n) \leftrightarrow \psi(n)) \rightarrow ([\varphi(0) \wedge \forall n(\varphi(n) \rightarrow \varphi(n+1))] \rightarrow \forall n \varphi(n)),$$

where φ is Σ^0_n and ψ is Π^0_n.

A summary of basic equivalences

Let $n \geq 1$. Over RCA_0:

- $\mathsf{I}\Sigma^0_n \leftrightarrow \mathsf{I}\Pi^0_n$;
- $\mathsf{B}\Sigma^0_{n+1} \leftrightarrow \mathsf{B}\Pi^0_n$;
- $\mathsf{I}\Sigma^0_{n+1} \Rightarrow \mathrm{B}\Sigma^0_{n+1} \Rightarrow \mathsf{I}\Sigma^0_n$ (Kirby & Paris);
- if $n \geq 2$, then $\mathsf{I}\Delta^0_n \leftrightarrow \mathsf{B}\Sigma^0_n$ (Slaman).

Two points:

(1) RCA_0 proves $\mathsf{I}\Pi^0_1$.

(2) The bounding axioms are equivalent to induction axioms.

Contents

Equivalence between a theorem and an induction scheme

$\mathsf{RT}^1_{<\mathbb{N}}$ is the following statement (it's the infinite pigeonhole principle):

For every $k > 0$ and every $f\colon \mathbb{N} \to k$, there is an infinite $H \subseteq \mathbb{N}$ and a $c < k$ such that $(\forall n \in H)(f(n) = c)$.

Theorem (Hirst)

$\mathsf{B}\Sigma^0_2$ *and* $\mathsf{RT}^1_{<\mathbb{N}}$ *are equivalent over* RCA_0*.*

Remember that $\mathsf{B}\Sigma^0_2 \leftrightarrow \mathsf{B}\Pi^0_1$ over RCA_0 (from the previous slide).

So Hirst's theorem also means that $\mathsf{B}\Pi^0_1$ and $\mathsf{RT}^1_{<\mathbb{N}}$ are equivalent over RCA_0.

Let $f\colon \mathbb{N} \to k$.

Suppose for a contradiction that no color $c < k$ appears infinitely often:

$$(\forall c < k)(\exists n)(\forall m)(m > n \to f(m) \neq c).$$

Then by $\mathsf{B}\Pi^0_1$ there is a b such that

$$(\forall c < k)(\exists n < b)(\forall m)(m > n \to f(m) \neq c).$$

This means that $f(b+1) \not< k$, a contradiction. So in fact some color $c < k$ appears infinitely often.

Let $H = \{n : f(n) = c\}$.

$\mathsf{RCA}_0 + \mathsf{RT}^1_{<\mathbb{N}} \vdash \mathsf{B}\Pi^0_1$

Suppose

$$(\forall n < t)(\exists m)(\forall z)\varphi(n, m, z)$$

(where φ has only bounded quantifiers).

Define $f\colon \mathbb{N} \to \mathbb{N}$ by

$$f(\ell) = \begin{cases} (\mu b < \ell)(\forall n < t)(\exists m < b)(\forall z < \ell)\varphi(n, m, z) & \text{if a } b \text{ exists} \\ \ell & \text{otherwise} \end{cases}$$

$f(\ell)$ is the least b that witnesses bounding when only looking up to ℓ.

$\mathsf{RCA}_0 + \mathsf{RT}^1_{<\mathbb{N}} \vdash \mathsf{B}\Pi^0_1$

Reminder:

$$f(\ell) = \begin{cases} (\mu b < \ell)(\forall n < t)(\exists m < b)(\forall z < \ell)\varphi(n, m, z) & \text{if a } b \text{ exists} \\ \ell & \text{otherwise} \end{cases}$$

If $\operatorname{ran}(f)$ is bounded, then $\mathsf{RT}^1_{<\mathbb{N}}$ applies, and there is a b and an infinite H such that $(\forall \ell \in H)(f(\ell) = b)$.

In this case, one proves that b is the desired bound:

$$(\forall n < t)(\exists m < b)(\forall z)\varphi(n, m, z)$$

Otherwise $\operatorname{ran}(f)$ is unbounded. Define a sequence $(\ell_i : i \in \mathbb{N})$ so that $\ell_i < \ell_{i+1}$ and $f(\ell_i) < f(\ell_{i+1})$.

Now define $g \colon \mathbb{N} \to t$ by

$$g(i) = (\mu n < t)(\forall m < f(\ell_i) - 1)(\exists z < \ell_i)(\neg\varphi(n, m, z)).$$

$\mathsf{RCA}_0 + \mathsf{RT}^1_{<\mathbb{N}} \vdash \mathsf{B}\Pi^0_1$

Reminder:

$$g(i) = (\mu n < t)(\forall m < f(\ell_i) - 1)(\exists z < \ell_i)(\neg\varphi(n, m, z)).$$

By $\mathsf{RT}^1_{<\mathbb{N}}$ applied to g, there is an $n < t$ and an infinite H such that $(\forall i \in H)(g(i) = n)$.

Let m be such that $\forall z \varphi(n, m, z)$.

Let $i \in H$ be such that $f(\ell_i) - 1 > m$.

Then $g(i) = n$, so $\exists z(\neg\varphi(n, m, z))$. Contradiction!

Contents

Π^1_1-conservativity

Definition (Π^1_1-conservativity)

Let S and T be theories with $\mathsf{S} \subseteq \mathsf{T}$. Then T is Π^1_1-conservative over S if $\varphi \in \mathsf{T} \leftrightarrow \varphi \in \mathsf{S}$ whenever φ is a Π^1_1 sentence.

So although T may be stronger than S, this additional strength is not witnessed by a Π^1_1 sentence.

(Of course you can study conservativity for other formula classes too.)

Classic examples:

- WKL_0 is Π^1_1-conservative over RCA_0 (Harrington).
- $\mathsf{WKL}_0 + \mathsf{B}\Sigma^0_2$ is Π^1_1-conservative over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$ (Hájek).

Why study conservativity?

Conservativity gives a way to express that T is stronger than S but not too much stronger than S.

Conservativity is useful for studying the first-order consequences of a theory.

For example, RCA_0 and WKL_0 have the same first-order part because WKL_0 is Π^1_1-conservative over RCA_0.

How do you prove conservativity results?

Proving that T is Π^1_1-conservative over S is about proving that every countable model of S can be extended to a model of T with the same first-order part.

For example:

Theorem (Harrington)

Every countable model of RCA_0 *is a submodel of a countable model of* WKL_0 *with the same first-order part.*

How do you prove conservativity results?

Suppose $\mathsf{RCA}_0 \nvdash \forall X \varphi(X)$, where φ is arithmetic.

Then there is a countable model $M = (\mathbb{N}, \mathcal{S})$ such that $M \vDash \mathsf{RCA}_0$ and $M \vDash \exists X(\neg\varphi(X))$.

Let $X \in \mathcal{S}$ witness $M \vDash \neg\varphi(X)$.

M is a submodel of a countable model $N = (\mathbb{N}, \mathcal{T})$ of WKL_0, where $\mathcal{S} \subseteq \mathcal{T}$.

Thus $X \in \mathcal{T}$ and so $N \vDash \neg\varphi(X)$. This is because φ is arithmetic, so the truth of $\varphi(X)$ depends only on $\mathbb{N}$ and X.

So $N \vDash \mathsf{WKL}_0$ and $N \vDash \exists X(\neg\varphi(X))$.

So $\mathsf{WKL}_0 \nvdash \forall X \varphi(X)$.

Extending models while preserving induction

Suppose φ and ψ are two theorems, and we want to prove that $\mathsf{RCA}_0 + \varphi \nvdash \psi$.

So we have to build a model of $\mathsf{RCA}_0 + \varphi$ that is not a model of ψ.

In many cases, this can be done by working over ω and proving that you can add sets witnessing φ that do not compute sets witnessing ψ.

Induction is never a problem. With ω, you have as much induction as you could ever want, and you can never add a set that breaks $\mathsf{I}\Sigma^0_1$.

When extending models, the situation is different. You are given a model $(\mathbb{N}, \mathcal{S})$ of RCA_0, and all you know is that $\mathsf{I}\Sigma^0_1$ holds relative to the $X \in \mathcal{S}$.

When adding sets to $\mathcal{S}$, you need to make sure that $\mathsf{I}\Sigma^0_1$ holds relative to them.

Weak Ramsey principles

Definition (RT^2_2, ADS, and CAC)

- RT^2_2: For every $f\colon [\mathbb{N}]^2 \to 2$, there is an infinite H such that $|f([H]^2)| = 1$.
- CAC: In every partial order, there is either an infinite chain or an infinite antichain.
- ADS: In every linear order, there is either an infinite increasing sequence or an infinite decreasing sequence.

RT^2_2 is strictly stronger than CAC over RCA_0 (Hirschfeldt & Shore).
CAC is strictly stronger than ADS over RCA_0 (Lerman, Solomon, Towsner).

RT^2_2 proves $\mathsf{B}\Sigma^0_2$ over RCA_0 (Hirst).
CAC and ADS both prove $\mathsf{B}\Sigma^0_2$ over RCA_0 (Chong, Lempp, Yang).

What about $\mathsf{I}\Sigma^0_2$?

Question

Does RT^2_2 *imply* $\mathsf{I}\Sigma^0_2$ *over* RCA_0*? Does* CAC*? Does* ADS*?*

Answer: None of these principles implies $\mathsf{I}\Sigma^0_2$ over RCA_0 (Chong, Slaman, Yang).

Theorem (Chong, Slaman, Yang)

Both ADS *and* CAC *are* Π^1_1*-conservative over* $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$.

It follows that neither ADS nor CAC proves $\mathsf{I}\Sigma^0_2$ over RCA_0 because $\mathsf{I}\Sigma^0_2$ consists of arithmetical axioms, and $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2 \nvdash \mathsf{I}\Sigma^0_2$.

Contents

The pigeonhole principle for trees

Definition (TT^1)

TT^1: For every $k > 0$ and every $f\colon 2^{<\mathbb{N}} \to k$, there is an infinite $T \subseteq 2^{<\mathbb{N}}$ that is order-isomorphic to $2^{<\mathbb{N}}$ and a $c < k$ such that $(\forall \sigma \in T)(f(\sigma) = c)$.

Theorem

- $\mathsf{RCA}_0 + \mathsf{I}\Sigma^0_2 \vdash \mathsf{TT}^1$ (Chubb, Hirst, McNicholl).
- $\mathsf{RCA}_0 + \mathsf{TT}^1 \vdash \mathsf{B}\Sigma^0_2$ (Chubb, Hirst, McNicholl).
- $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2 \nvdash \mathsf{TT}^1$ (Corduan, Groszek, Mileti).

Theorem (Breaking news! Chong and Li)

$\mathsf{RCA}_0 + \mathsf{TT}^1 \vdash \mathsf{I}\Sigma^0_2$

$\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2 \nvdash \mathsf{TT}^1$

Theorem (Corduan, Groszek, Mileti)

If S *is an extension of* RCA_0 *by* Π^1_1 *axioms, then* $\mathsf{S} \vdash \mathsf{TT}^1$ *if and only if* $\mathsf{S} \vdash \mathsf{I}\Sigma^0_2$.

Thus $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2 \nvdash \mathsf{TT}^1$ because $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$ is strictly weaker than $\mathsf{RCA}_0 + \mathsf{I}\Sigma^0_2$.

The idea here is to prove that if $M \models \mathsf{RCA}_0$ but $M \not\models \mathsf{I}\Sigma^0_2$, then there is a coloring $f\colon 2^{<\mathbb{N}} \to k$ such that no $T \leq_{\mathsf{T}} f$ is a monochromatic subtree of $2^{<\mathbb{N}}$ isomorphic to $2^{<\mathbb{N}}$.

Thus if $M \models \mathsf{RCA}_0$ and $M \not\models \mathsf{I}\Sigma^0_2$, then there is a submodel N of M with the same first-order part that satisfies all the Π^1_1 sentences true in M but not TT^1.

Contents

Diagonally non-recursive functions

Definition

Let $f, g\colon \mathbb{N} \to \mathbb{N}$, and let $k \geq 2$.

- f is diagonally non-recursive relative to g (i.e., $\mathsf{DNR}(g)$) if $\forall e(f(e) \neq \Phi_e^g(e))$.
- f is k-bounded diagonally non-recursive relative to g (i.e., $\mathsf{DNR}(k, g)$) if f is $\mathsf{DNR}(g)$ and $\operatorname{ran}(f) \subseteq \{0, 1, \ldots, k-1\}$.

(In the case $g = 0$, we just write DNR and $\mathsf{DNR}(k)$.)

Notice that every $\mathsf{DNR}(k)$ function is also a $\mathsf{DNR}(k+1)$ function.

Does every $\mathsf{DNR}(k+1)$ function compute a $\mathsf{DNR}(k)$ function?

Classical theorems concerning DNR(k) functions

Theorem (Jockusch attributes to Friedberg)

For every $k \geq 2$, *every* DNR(k) *function computes some* DNR(2) *function.*

Theorem (Jockusch)

The reduction from DNR(k) *to* DNR(2) *is not uniform. That is, for each* $k \geq 2$, *there is no Turing functional* Φ *such that* $(\forall f \in \mathsf{DNR}(k+1))(\Phi^f \in \mathsf{DNR}(k))$.

Theorem (Jockusch & Soare)

Every DNR(2) *function computes an infinite path through every infinite, computable tree* $T \subseteq 2^{<\mathbb{N}}$.

Reverse math and DNR functions

Here we overload the 'DNR' notation:

- $\mathsf{DNR}(g)$ also denotes the statement "there is an f that is $\mathsf{DNR}(g)$."
- $\mathsf{DNR}(k, g)$ also denotes the statement "there is an f that is $\mathsf{DNR}(k, g)$."

Theorem

- $\mathsf{RCA}_0 \vdash \mathsf{WKL} \leftrightarrow \forall g(\mathsf{DNR}(2, g))$.
- *For each standard* $k \geq 2$, $\mathsf{RCA}_0 \vdash \mathsf{WKL} \leftrightarrow \forall g(\mathsf{DNR}(k, g))$.

Theorem (Ambos-Spies, Kjos-Hanssen, Lempp, Slaman)

$\forall g(\mathsf{DNR}(g))$ *is strictly weaker than* WWKL *over* RCA_0.

$\exists k \forall g(\mathsf{DNR}(k,g))$ versus WKL

What if you know that there are $\mathsf{DNR}(k,g)$ functions for some $k \geq 2$, but you do not know which k?

This is essentially a question of Simpson, who asked:

Question (Simpson, 2001)

Is $\exists k \forall g(\mathsf{DNR}(k,g))$ *equivalent to* WKL *over* RCA_0*?*

Of course, we could ask the same question about the (formally weaker) statement $\forall g \exists k(\mathsf{DNR}(k,g))$.

Connection to graph colorability

By compactness, every locally ℓ-colorable graph is ℓ-colorable.

There has been considerable work analyzing the computability-theoretic content of this fact, culminating (for our purposes) in a theorem of Schmerl:

Theorem (Schmerl)

Fix standard ℓ and k such that $2 \leq \ell \leq k$. Then the following are equivalent over RCA_0*:*

(i) WKL

(ii) *Every locally ℓ-colorable graph is k-colorable.*

The content of this theorem is that the ability to color a graph with a fixed sub-optimal number of colors is as strong as the ability to color the graph with the optimal number of colors.

Connection to graph colorability

Denote by COL(ℓ, k, G) the statement "if G is a locally ℓ-colorable graph, then G is k-colorable."

Question (essentially of Simpson)

Are either of the following statements equivalent to WKL *over* RCA_0*?*

- $\forall \ell \exists k \forall G(\mathsf{COL}(\ell, k, G))$
- $\forall \ell \forall G \exists k(\mathsf{COL}(\ell, k, G))$

DNR(k) and induction

Recall: every DNR(k) function computes a DNR(2) function.

With a little work, the proof can be implemented using $\mathsf{I}\Sigma^0_2$.

Hence the following are equivalent over $\mathsf{RCA}_0 + \mathsf{I}\Sigma^0_2$:

(i) WKL

(ii) $\exists k \forall g(\mathsf{DNR}(k, g))$

(iii) $\forall g \exists k(\mathsf{DNR}(k, g))$

Dorais, Hirst, and I show that the three equivalences fail if you only assume $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$.

The strength of DNR(k) functions

We show that, over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$,

$$\mathsf{WKL} \Rightarrow \exists k \forall g(\mathsf{DNR}(k,g)) \Rightarrow \forall g \exists k(\mathsf{DNR}(k,g)), \text{ and}$$
$$\mathsf{WKL} \Rightarrow \forall \ell \exists k \forall G(\mathsf{COL}(\ell,k,G)), \text{ and}$$
$$\forall \ell \forall G \exists k(\mathsf{COL}(\ell,k,G)) \not\Rightarrow \exists k \forall g(\mathsf{DNR}(k,g))$$

The results rewritten in an authoritative purple box:

Theorem (Dorais, Hirst, S)

(i) $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2 + \exists k \forall g(\mathsf{DNR}(k,g)) \nvdash \mathsf{WKL}$

(ii) $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2 + \forall g \exists k(\mathsf{DNR}(k,g)) \nvdash \exists k \forall g(\mathsf{DNR}(k,g))$

(iii) $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2 + \forall \ell \exists k \forall G(\mathsf{COL}(\ell,k,G)) \nvdash \mathsf{WKL}$

(iv) $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2 + \forall \ell \forall G \exists k(\mathsf{COL}(\ell,k,G)) \nvdash \exists k \forall g(\mathsf{DNR}(k,g))$

Remaining questions

Are $\exists k \forall g(\mathsf{DNR}(k,g))$ and $\exists k \forall G(\mathsf{COL}(k,G))$ equivalent over RCA_0?

Are $\forall g \exists k(\mathsf{DNR}(k,g))$ and $\forall \ell \forall G \exists k(\mathsf{COL}(\ell,k,G))$ equivalent over RCA_0?

Does $\forall \ell \exists k \forall G(\mathsf{COL}(\ell,k,G))$ (or $\forall \ell \forall G \exists k(\mathsf{COL}(\ell,k,G))$) imply WKL over $\mathsf{RCA}_0 + \mathsf{I}\Sigma^0_2$?

Is $\forall \ell \forall G \exists k(\mathsf{COL}(\ell,k,G))$ strictly weaker than $\forall \ell \exists k \forall G(\mathsf{COL}(\ell,k,G))$ over RCA_0 (or over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$)?

Best question:
Does $\exists k \forall g(\mathsf{DNR}(k,g))$ imply WWKL over RCA_0 (or over $\mathsf{RCA}_0 + \mathsf{B}\Sigma^0_2$)?

Thank you!

Thank you for coming to my talk!
Do you have a question about it?