Reverse Mathematics and Computable Analysis

Talk based on a paper written with Guido Gherardi

Alberto Marcone

Dipartimento di Scienze Matematiche, Informatiche e Fisiche
Università di Udine
Italy
alberto.marcone@uniud.it
http://www.dimi.uniud.it/marcone

XXVI incontro dell'AILA
September 25–28, 2017

Mathematicians make statements of the form:

- Theorems Φ and Ψ are equivalent;
- Theorem Φ is (properly) stronger than theorem Ψ.

Taken literally, the first statement is trivially true, and the second one is trivially false.

But we know what these assertions mean.

In the last decades, mathematical logic has tried to give a rigorous meaning to statements of this kind. Reverse mathematics and computable analysis are two programs addressing this topic.

The two approaches were brought together in the 2015 workshop I co-organized in Dagstuhl.

Outline

1. **Reverse mathematics**

2. **Computable analysis**

3. **Analogies and differences between the two approaches**

Equivalences

> *Reciprocation of premisses and conclusion is more frequent in mathematics [than in] dialectical disputations*
> *(Aristotle, Posterior Analytics, 78a10)*

Aristotle had in mind equivalences such as "a triangle has two congruent sides if and only if it has two congruent angles".

On the other hands, modern mathematics includes many equivalences between statements, such as the one between the axiom of choice and Zorn's lemma on the basis of ZF.

All results of this sort can be dubbed "reverse mathematics".

However the term usually applies to research carried out in the context of subsystems of second order arithmetic.

Second order arithmetic

The language of second order arithmetic $\mathcal{L}_2$ has two sorts of variables:
one for natural numbers, the other for sets of natural numbers.

There are symbols for basic algebraic operations,
for equality between natural numbers,
and for membership between a number and a set.

We use classical logic.

Full second order arithmetic Z_2 is the theory with algebraic axioms for the natural numbers, full induction, and full comprehension:

$$\exists X \, \forall n \, (n \in X \iff \varphi(n)),$$

with X not free in $\varphi(n)$.

Semantics of second order arithmetic

A model for $\mathcal{L}_2$ has the form

$$M = (|M|, \mathcal{S}_M, 0_M, 1_M, +_M, \cdot_M, <_M)$$

where $|M|$ serves as the range of the number variables, $\mathcal{S}_M$ is a set of subsets of $|M|$ serving as the range of the set variables.

An ω-model is an $\mathcal{L}_2$ model M whose first order part is standard, i.e. of the form $(\omega, 0, 1, +, \cdot, <)$.
Thus M can be identified with the collection of sets of natural numbers serving as the range of the set variables in $\mathcal{L}_2$.

For example REC is the ω-model consisting of the computable (or recursive) sets.

Mathematics in second order arithmetic

The idea that in (subsystems of) second order arithmetic it is possible to state and prove many significant mathematical theorems goes back to Hermann Weyl, Hilbert and Bernays.

The systematic search for the subsystems of second order arithmetic which are sufficient and necessary to prove these theorems was started by Harvey Friedman around 1970, and pursued by Steve Simpson and many others.

Subsystems of second order arithmetic

1. RCA_0: algebraic axioms for the natural numbers, $\mathbf{\Sigma}^0_1$-induction, and comprehension for $\mathbf{\Delta}^0_1$ formulas ω^ω
2. $\mathsf{WKL}_0 = \mathsf{RCA}_0 +$ König's lemma for binary trees ω^ω
3. ACA_0: comprehension extended to arithmetical formulas ε_0
4. $\mathsf{ATR}_0 = \mathsf{ACA}_0 +$ defin. by arithmetical transfinite recursion Γ_0
5. $\mathbf{\Pi}^1_1$-CA_0: comprehension extended to $\mathbf{\Pi}^1_1$ formulas $\Psi_{\Omega_1}(\Omega_\omega)$

RCA_0 is the base theory for reverse mathematics: it allows the development of "computable mathematics".

RCA_0 and WKL_0 are $\mathbf{\Pi}^0_2$-conservative over PRA.

ACA_0 is conservative over PA.

Reverse mathematics results: RCA_0

RCA_0 proves the following statements:

1. The intersection of a sequence of intervals of $\mathbb{R}$, each included in the preceding one and with length going to 0, consists exactly of a real number.
2. A form of Baire category theorem.
3. The intermediate value theorem.
4. The Tietze extension theorem.
5. A weak form of Gödel's completeness theorem: every countable set of first-order formulas closed under logical consequence and consistent has a (countable) model.
6. Every countable field has an algebraic closure.
7. The uniform boundedness theorem for pointwise bounded sequences of operators on Banach spaces.

Reverse mathematics results: WKL_0

Over RCA_0, WKL_0 is equivalent to the following statements:

1. The Heine-Borel compactness theorem for $[0, 1]$.
2. Every continuous function from $[0, 1]$ to $\mathbb{R}$ is uniformly continuous.
3. The Cauchy-Peano local existence theorem for ordinary differential equations.
4. The Hahn-Banach theorem for separable Banach spaces.
5. Every torsion free abelian group is orderable.
6. Every countable commutative ring with identity has a prime ideal.
7. Every countable field has a unique algebraic closure.
8. Gödel's completeness theorem.

Reverse mathematics results: ACA_0

Over RCA_0, ACA_0 is equivalent to the following statements:

1. The Bolzano-Weierstraß theorem for bounded sequences of reals.
2. The Ascoli-Arzelà lemma about bounded and equicontinuous sequences of functions on bounded intervals in $\mathbb{R}$.
3. Every countable commutative ring with identity has a maximal ideal.
4. Hahn's theorem: every ordered abelian group can be embedded in a product of copies of $(\mathbb{R}, +)$.
5. König's lemma.
6. Ramsey theorem for exponent k when $k \geq 3$: for every $c : [\mathbb{N}]^k \to \{0, \dots \ell - 1\}$ there exists $H \subseteq \mathbb{N}$ infinite and such that $c \restriction [H]^k$ is constant. RT^k_ℓ

Reverse mathematics results: ATR_0

Over RCA_0, ATR_0 is equivalent to the following statements:

1. The perfect set theorem for uncountable closed subsets of $\mathbb{R}$.
2. Lusin separation theorem.
3. The domain of a single-valued Borel set in the plane is Borel.
4. Determinacy of open games in $\mathbb{N}^{\mathbb{N}}$.
5. The Galvin-Prikry theorem for open subsets of $[\mathbb{N}]^{\mathbb{N}}$.
6. Ulm classification theorem about abelian p-groups.
7. Given two countable well orders, one of them embeds into the other.
8. Hausdorff classification of countable scattered linear orders.

Reverse mathematics results: $\mathbf{\Pi}^1_1$-CA_0

Over RCA_0, $\mathbf{\Pi}^1_1$-CA_0 is equivalent to the following statements:

1. Cantor-Bendixson theorem about closed subsets of $\mathbb{R}$.
2. Silver's theorem about co-analytic equivalence relations on $\mathbb{R}$.
3. Determinacy of games with payoff the intersection of an open and a closed set in $\mathbb{N}^{\mathbb{N}}$.
4. The Galvin-Prikry theorem for subsets of $[\mathbb{N}]^{\mathbb{N}}$ of finite Borel rank.
5. Every countable abelian group is the direct sum of a divisible group and a reduced group.
6. Mal'tsev theorem about the order type of countable ordered groups.

The big five of reverse mathematics

RCA_0, WKL_0, ACA_0, ATR_0, and $\mathbf{\Pi}^1_1$-CA_0 have been claimed to correspond also to different approaches to the foundations of mathematics.

They can also be viewed as assertions about the existence of more and more incomputable sets and so connected to basic theorems from computability theory.

The wealth of results showing their equivalence with mathematical theorems led to the terminology the big five.

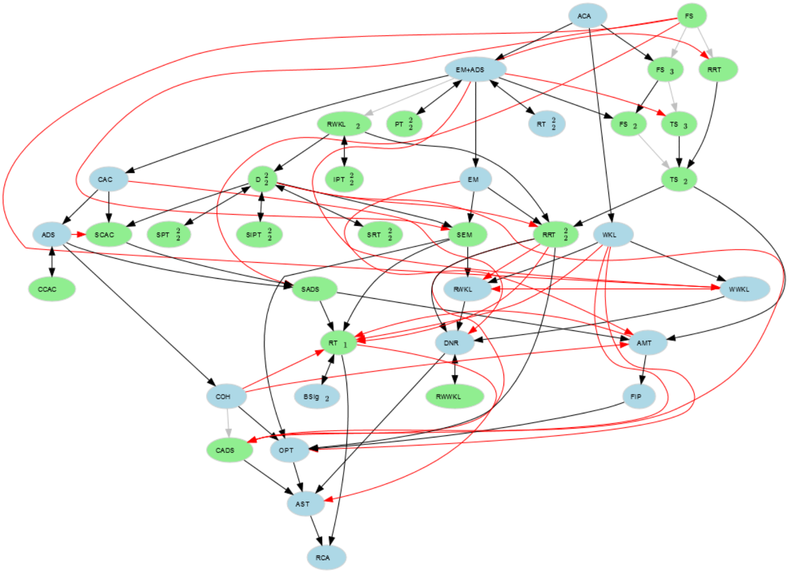
The zoo

In 1995 Seetapun showed that Ramsey Theorem for pairs RT^2_2 does not imply ACA_0.
It was already known that WKL_0 does not prove RT^2_2.
In 2012 Liu showed that RT^2_2 does not imply WKL_0.

After Seetapun's result, many statements provable in ACA_0, unprovable in RCA_0, and not equivalent to either ACA_0 or WKL_0, have been discovered.

The neat five-levels building of XXth century reverse mathematics is now much more complex, with lots of different beasts:
the zoo of XXIst century reverse mathematics.

A picture of a portion of the zoo (from Ludovic Patey's website)



Three beasts in the zoo

RT^2_2 Ramsey Theorem for Pairs: for every
$f : [\mathbb{N}]^2 \to \{0,1\}$
there exist $i < 2$ and $H \subseteq \mathbb{N}$ infinite
such that $f(n,m) = i$ for every $n,m \in H$

CAC Chain-Antichain: every infinite partial order has either an infinite chain or an infinite antichain

ADS Ascending Sequence-Descending Sequence: every infinite linear order has either an infinite ascending sequence or an infinite descending sequence

$$\begin{tikzcd}
\mathsf{ACA}_0 \arrow[rr, Rightarrow] \arrow[dr, Rightarrow] & & \mathsf{WKL}_0 \arrow[rr, Rightarrow] & & \mathsf{RCA}_0 \\
& \mathsf{RT}^2_2 \arrow[r, Rightarrow] & \mathsf{CAC} \arrow[r, Rightarrow] & \mathsf{ADS} \arrow[ur, Rightarrow] &
\end{tikzcd}$$

A mutation in reverse mathematics

The original formulation of the reverse mathematics program highlighted provability, and often reverse mathematics was viewed as part of proof theory.

Nowadays the typical proof of a result in the zoo use computability theory constructions based on forcing or hyperimmunity arguments.

These arguments are used to build ω-models of one statement but not of the other.

These leads to concentrate more on logical consequence for a semantics where interpretations are ω-models.

Shore and others put forward the idea of studying the web of implications and non-implications considering only ω-models.

Some Italian contributors to reverse mathematics

- Lorenzo Carlucci
- Emanuele Frittaion
- Mariagnese Giusto
- Alberto Marcone
- Silvia Steila

Computable analysis

1. Reverse mathematics

2. **Computable analysis**

3. Analogies and differences between the two approaches

Extending classical computability

Classical computability deals with functions and relation on the set of the natural numbers or other countable sets.

Several nonequivalent approaches to computability theory for the reals have been proposed in the literature.

Type-2 Theory of Effectivity (TTE) extends the ordinary notion of Turing computability to second countable T_0-topological spaces, and therefore deals with computability over the reals as a particular case within a more general theory of computability with mathematical objects.

Computing with infinite objects

Concrete computing machines do not manipulate directly mathematical objects, but they perform computations on sequences of digits which are codings for such objects.

Most mathematical objects require an infinite amount of information to be completely described, and we need to extend the classical theory of computation to infinite sequences.

We assume that an infinite object is described by giving better and better approximations and that the output of the computation is also produced by approximations that become more precise as the computation proceeds.

Thus we expect that the computation never halts.

The abstract model of computation

TTE Turing machines have one input tape, one working tape and one output tape. Each tape has a head and can store a natural number in every cell.

All ordinary instructions for Turing machines are allowed for the working tape, while the head of the input tape can only read, and the head of the output tape can only write and move forward.

In a computation we supply an element of $\mathbb{N}^{\mathbb{N}}$ on the input tape and the machine will write an element of $\mathbb{N}^{\mathbb{N}}$ on the output tape.

Since we do not correct the output each partial output is reliable.

TTE Turing machines can also be realized as ordinary oracle Turing machines: the oracle supplies the input and the n-th bit of the output is computed by letting the oracle Turing machine work on n.

TTE computability

The partial functions from $\mathbb{N}^\mathbb{N}$ to $\mathbb{N}^\mathbb{N}$ computed by TTE machines are the computable partial functions from $\mathbb{N}^\mathbb{N}$ to $\mathbb{N}^\mathbb{N}$ (aka Lachlan functionals).

Notice that every computable partial function is continuous.

So far we are only computing functions from $\mathbb{N}^\mathbb{N}$ to $\mathbb{N}^\mathbb{N}$.

Represented spaces

A representation σ_X of a set X is a surjective partial function $\sigma_X : \subseteq \mathbb{N}^\mathbb{N} \to X$.

The pair (X, σ_X) is a represented space.

If $x \in X$ a σ_X-name for x is any $p \in \mathbb{N}^\mathbb{N}$ such that $\sigma_X(p) = x$.

$x \in X$ is computable (w.r.t. σ_X) if it has a computable σ_X-name.

Representations are analogous to the codings used in reverse mathematics to speak about various mathematical objects in subsystems of second order arithmetic.

Two examples of representations

Let (X, D, d) be a computable metric space.

Cauchy representation of X: $p \in \mathbb{N}^{\mathbb{N}}$ is a name for $x \in X$ if p lists a sequence $x_i \subseteq D$ such that $d(x_i, x_{i+1}) \leq 2^{-i}$ for every i and $\lim x_i = x$.

negative representation of the set $\mathcal{A}^{-}(X)$ of closed subsets of X: $p \in \mathbb{N}^{\mathbb{N}}$ is a name for the closed set C if p lists a sequence of open balls with center in D and rational radius whose union is $X \setminus C$.

If $X = \mathbb{N}^{\mathbb{N}}$ the Cauchy representation of x essentially means giving a digit of x at every step, and the negative representation is computably equivalent to the representation of C by (the characteristic function of) a tree $T \subseteq \mathbb{N}^{<\mathbb{N}}$ such that $[T] = C$.

Computable functions between represented spaces

If (X, σ_X) and (Y, σ_Y) are represented spaces and $f :\subseteq X \rightrightarrows Y$ we say that f is computable if there exists a computable $F :\subseteq \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ such that $\sigma_Y(F(p)) \in f(\sigma_X(p))$ whenever $f(\sigma_X(p))$ is defined, i.e. p is a name for an element of $\mathrm{dom}(f)$.

Such an F is called a realizer of f.

Extending reducibilities

Reducibilities such as many-one, Turing or polynomial-time have been an extraordinarily important tool in computability theory and theoretical computer science from their very beginnings.

In recent years these reducibilities have been transferred to the continuous setting, where they allow us to classify computational problems on real numbers and other continuous data types.

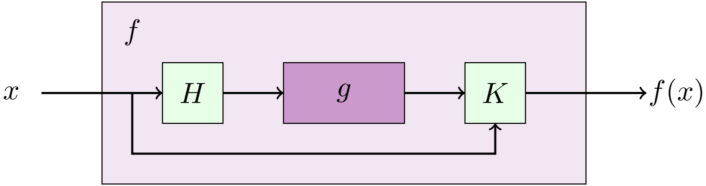
In the late 1980s Klaus Weihrauch introduced a reducibility that can be seen as an analogue of many-one reducibility for (multi-valued) functions on infinite data types.

This reducibility, now called Weihrauch reducibility, was studied since the 1990s by Weihrauch's school of computable analysis and starting from Gherardi-M 2009 it is used as a tool for comparing the strength of mathematical statements.

Weihrauch reducibility

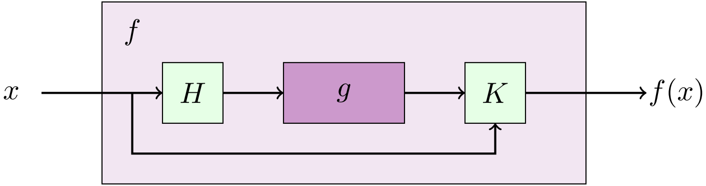
Let $f :\subseteq X \rightrightarrows Y$ and $g :\subseteq Z \rightrightarrows W$ be partial multi-valued functions between represented spaces.

f is Weihrauch reducible to g, $f \leq_{\mathrm{W}} g$, if there are computable $H :\subseteq X \rightrightarrows Z$ and $K :\subseteq X \times W \rightrightarrows Y$ such that $K(x, gH(x)) \subseteq f(x)$ for all $x \in \mathrm{dom}(f)$:



In other words, for all $x \in \mathrm{dom}(f)$, we have $H(x) \subseteq \mathrm{dom}(g)$ and $K(x, w) \subseteq f(x)$ for every $w \in g(H(x))$.

Weihrauch reducibility



$f \leq_{\mathrm{W}} g$ means that the problem of computing f can be computably and uniformly solved by using in each instance a single computation of g:

H modifies the input of f to feed it to g, while K, using also the original input, transforms the output of g into the correct output of f.

Another characterization of Weihrauch reducibility is provided by the fact that $f \leq_{\mathrm{W}} g$ if and only if there is a Turing machine that computes (a realizer of) f using (a realizer of) g as an oracle exactly once during its infinite computation.

The Weihrauch lattice

$\leq_W$ is reflexive and transitive and induces the equivalence relation $\equiv_W$.

The $\equiv_W$-equivalence classes are called Weihrauch degrees.

The partial order on the sets of Weihrauch degrees is a distributive bounded lattice with several natural and useful algebraic operations: the Weihrauch lattice.

The Weihrauch lattice and mathematical practice

The Weihrauch lattice allows a calculus of mathematical problems.

A mathematical problem can be identified with a partial multi-valued function $f :\subseteq X \rightrightarrows Y$: there are sets of potential inputs X and outputs Y, $\mathrm{dom}(f) \subseteq X$ contains the valid instances of the problem, and $f(x)$ is the set of solutions of the problem f for instance x.

If X and Y are represented spaces and $\forall x \in X(\varphi(x) \rightarrow \exists y \in Y\ \psi(x,y))$ is a true statement, we consider the mathematical problem with domain $\{\, x \in X \mid \varphi(x) \,\}$ such that $f(x) = \{\, y \in Y \mid \psi(x,y) \,\}$.

The Zero Problem for a topological space X is the partial multi-valued function $\mathsf{Z}_X :\subseteq C(X, \mathbb{R}) \rightrightarrows X$ defined by $\mathsf{Z}_X(h) = \{\, x \in X \mid h(x) = 0 \,\}$;
$\mathrm{dom}(\mathsf{Z}_X)$ is the set of continuous functions with at least one zero.

Choice functions and parallelization

Let X be a computable metric space and $\mathcal{A}^-(X)$ be the space of its closed subsets represented by negative information.

$\mathsf{C}_X :\subseteq \mathcal{A}^-(X) \rightrightarrows X$ is the choice function for X: it picks from a nonempty closed set in X one of its elements.

It is easy to see that $\mathsf{C}_X \equiv_{\mathrm{W}} \mathsf{Z}_X$ and that already C_2 is noncomputable.

$\mathsf{UC}_X :\subseteq \mathcal{A}^-(X) \to X$ is the unique choice function for X: it picks from a singleton (represented as a closed set) in X its unique element.

UC_2 is computable and, for example, $\mathsf{UC}_{\mathbb{N}} \equiv_{\mathrm{W}} \mathsf{UC}_{\mathbb{R}} \equiv_{\mathrm{W}} \mathsf{C}_{\mathbb{N}}$.

If $f :\subseteq X \rightrightarrows Y$ is a multi-valued function, the parallelization of f is the multi-valued function $\widehat{f} : X^{\mathbb{N}} \rightrightarrows Y^{\mathbb{N}}$ with $\mathrm{dom}(\widehat{f}) = \mathrm{dom}(f)^{\mathbb{N}}$ defined by $f((x_n)_{n\in\mathbb{N}}) = \prod_{n\in\mathbb{N}} f(x_n)$.

$\widehat{f}$ computes f countably many times in parallel.

A sample of results 1

- The multi-valued functions arising from the following theorems are computable: Urysohn Extension Lemma, Urysohn-Tietze Extension Lemma, Heine-Borel Theorem, Weierstraß Approximation Theorem, Baire Category Theorem, ...
- The multi-valued functions arising from the following theorems are Weihrauch equivalent to $\mathrm{C}_{\mathbb{N}}$: Banach Inverse Mapping Theorem, Open Mapping Theorem, Closed Graph Theorem, Uniform Boundedness Theorem, (contrapositive of) Baire Category Theorem, ...
- The multi-valued functions arising from the following theorems are Weihrauch equivalent to $\mathrm{C}_{2^{\mathbb{N}}}$: Weak Kőnig Lemma, Hahn-Banach Extension Theorem, Brouwer Fixed Point Theorem, ...

A sample of results 2

- The multi-valued functions arising from the following theorems are Weihrauch equivalent to $\widehat{\mathsf{C}_{\mathbb{N}}}$: limit, Monotone Convergence Theorem, Radon-Nikodym Theorem, ...
- The multi-valued functions arising from the following theorems are Weihrauch equivalent to $\mathsf{UC}_{\mathbb{N}^{\mathbb{N}}}$: Comparability of Well-Orders, $\mathbf{\Delta}^0_1$-determinacy, ...
- The multi-valued functions arising from the following theorems are Weihrauch equivalent to $\mathsf{C}_{\mathbb{N}^{\mathbb{N}}}$: the Perfect Tree Theorem, ...

Some Italian contributors to computable analysis

- Guido Gherardi
- Stefano Galatolo
- Alberto Marcone

Analogies and differences between the two approaches

The Weihrauch lattice and reverse mathematics

In most cases the Weihrauch lattice refines the classification provided by reverse mathematics: statements which are equivalent over RCA_0 may give rise to functions with different Weihrauch degrees.

Weihrauch reducibility is finer because requires both uniformity and use of a single instance of the harder problem.

Moreover, although sometimes a mathematical theorem naturally leads to a single function to be studied in the Weihrauch lattice, in other cases several functions arise from the same theorem.

For example the Baire Category Theorem and its contrapositive are reverse mathematically equivalent, yet give rise to functions which are not Weihrauch equivalent.

An example: detecting iteration

Let RT^n_k be the infinite Ramsey theorem for n-tuples and k colors ($n \geq 1$ and $k \geq 2$).
We denote by RT^n_k also the multi-valued function arising from it.

In reverse mathematics, the strength of RT^n_k depends only on n.

Fix n: the obvious proof that RT^n_j implies RT^n_k for $k > j$ uses multiple applications of RT^n_j.

Theorem (Hirschfeldt-Jockusch 2016, Patey 2016, Rakotoniaina 2015)

If $n \geq 1$ and $k > j \geq 2$ then $\mathsf{RT}^n_k \not\leq_{\mathrm{W}} \mathsf{RT}^n_j$.

An example: lack of uniformity

The intermediate value theorem is provable in the base system RCA_0 by the standard proof: given $f : [0,1] \to \mathbb{R}$ is continuous and such that $f(0) \cdot f(1) < 0$ then either there exists $q \in \mathbb{Q}$ such that $f(q) = 0$ (and we are done), or we can run the bisection method to find a zero for f.

This proof is non-uniform.

Let IVT be the multi-valued function arising from the intermediate value theorem.

Theorem (Weihrauch 2000)

IVT *is not computable.*

Exceptions

There are however exceptions to the phenomena described above: in some cases the reverse mathematics approach may detect differences that Weihrauch reducibility misses:

"the computable analyst is allowed to conduct an unbounded search for an object that is guaranteed to exist by (nonconstructive) mathematical knowledge, whereas the reverse mathematician has the burden of an existence proof with limited means" (Gherardi-M 2009).

An example: searching on the basis of mathematical knowledge

In reverse mathematics, the Heine-Borel compactness (every open cover has a finite subcover) of $[0, 1]$ is equivalent to WKL_0, and hence not computably true.

Weihrauch showed that the corresponding multi-valued functions is computable: given an open cover we can search for a finite subcover.
The compactness of $[0, 1]$ insures that the search will succeed.

From the reverse mathematics viewpoint, this algorithm can be defined in RCA_0, but the proof of its termination requires WKL_0.

The end

Thank you for your attention!