Archive for Mathematical Logic
https://doi.org/10.1007/s00153-018-0657-9

Truth, disjunction, and induction

Ali Enayat[1] · **Fedor Pakhomov**[2]

Received: 25 June 2018 / Accepted: 17 December 2018

Abstract
By a well-known result of Kotlarski et al. (1981), first-order Peano arithmetic PA can be conservatively extended to the theory $\mathsf{CT}^-[\mathsf{PA}]$ of a truth predicate satisfying compositional axioms, i.e., axioms stating that the truth predicate is correct on atomic formulae and commutes with all the propositional connectives and quantifiers. This result motivates the general question of determining natural axioms concerning the truth predicate that can be added to $\mathsf{CT}^-[\mathsf{PA}]$ while maintaining conservativity over PA. Our main result shows that conservativity fails even for the extension of $\mathsf{CT}^-[\mathsf{PA}]$ obtained by the seemingly weak axiom of disjunctive correctness DC that asserts that the truth predicate commutes with disjunctions of arbitrary finite size. In particular, $\mathsf{CT}^-[\mathsf{PA}] + \mathsf{DC}$ implies $\mathsf{Con}(\mathsf{PA})$. Our main result states that the theory $\mathsf{CT}^-[\mathsf{PA}] + \mathsf{DC}$ coincides with the theory $\mathsf{CT}_0[\mathsf{PA}]$ obtained by adding Δ_0-induction in the language with the truth predicate. This result strengthens earlier work by Kotlarski (1986) and Cieśliński (2010). For our proof we develop a new general form of Visser's theorem on non-existence of infinite descending chains of truth definitions and prove it by reduction to (Löb's version of) Gödel's second incompleteness theorem, rather than by using the Visser–Yablo paradox, as in Visser's original proof (1989).

Keywords Axiomatic truth · Compositional theory of truth · Conservativity

Mathematics Subject Classification 03F30

The work of Fedor Pakhomov is supported by the Russian Science Foundation under Grant 16-11-10252 and performed at Steklov Mathematical Institute of Russian Academy of Sciences. F. Pakhomov was selected as one of the Young Russian Mathematics award winners, and he would like to thank its sponsors and jury; his work on this particular paper was funded from another source. Sections 2 and 4 of this paper were contributed by A. Enayat; Section 3 was contributed by F. Pakhomov; Sections 1 and 5 were contributed jointly by the authors. The authors are grateful to the anonymous referees for their meticulous and valuable feedback.

✉ Ali Enayat
ali.enayat@gu.se

Extended author information available on the last page of the article

Published online: 04 February 2019

1 Introduction

By a theorem of Krajewski, Kotlarski, and Lachlan [12], every countable recursively saturated model $\mathcal{M}$ of PA (Peano Arithmetic) carries a 'full satisfaction class', i.e., there is a subset S of the universe M of $\mathcal{M}$ that 'decides' the truth/falsity of each sentence of arithmetic in the sense of $\mathcal{M}$—even sentences of nonstandard length—while obeying the usual recursive clauses guiding the behavior of a Tarskian satisfaction predicate. This remarkable theorem implies that theory $\mathsf{CT}^-[\mathsf{PA}]$ (compositional truth over PA with induction only for the language $\mathcal{L}_\mathsf{A}$ of arithmetic) is *conservative* over PA, i.e., if an $\mathcal{L}_\mathsf{A}$-sentence φ is provable in $\mathsf{CT}^-[\mathsf{PA}]$, then φ is already provable in PA. New proofs of this conservativity result were given by Visser and Enayat [5] using basic model theoretic ideas, and by Leigh [14] using proof theoretic tools; these new proofs make it clear that in the Krajewski–Kotlarski–Lachlan theorem the theory PA can be replaced by any 'base' theory that supports a modicum of coding machinery for handling elementary syntax.

On the other hand, it is well-known [9, Thm. 8.39 and Cor. 8.40] that the consistency of PA (and much more) is readily provable in the stronger theory CT[PA], which is the result of strengthening $\mathsf{CT}^-[\mathsf{PA}]$ with the scheme of induction over natural numbers for all $\mathcal{L}_{\mathsf{A}+\mathsf{T}}$-formulae, where $\mathcal{L}_{\mathsf{A}+\mathsf{T}} := \mathcal{L}_\mathsf{A} \cup \{\mathsf{T}(x)\}$.[1] Indeed, it is straightforward to demonstrate the consistency of PA within the subsystem $\mathsf{CT}_1[\mathsf{PA}]$ of CT[PA], where $\mathsf{CT}_n[\mathsf{PA}]$ is the subtheory of CT[PA] with the scheme of induction over natural numbers limited to $\mathcal{L}_{\mathsf{A}+\mathsf{T}}$-formulae that are at most of complexity Σ_n [16, Thm. 2.8].

The discussion above leaves open whether $\mathsf{CT}_0[\mathsf{PA}]$ is conservative over PA. Kotlarski [11] established that $\mathsf{CT}_0[\mathsf{PA}]$ is a subtheory of $\mathsf{CT}^-[\mathsf{PA}] + \mathsf{Ref}(\mathsf{PA})$, where Ref(PA) is the $\mathcal{L}_{\mathsf{A}+\mathsf{T}}$-sentence stating that "every first order consequence of PA is true". Recently Łełyk [15] demonstrated that the converse also holds, which immediately implies that $\mathsf{CT}_0[\mathsf{PA}]$ is not conservative over PA since Con(PA) is readily provable in $\mathsf{CT}^-[\mathsf{PA}] + \mathsf{Ref}(\mathsf{PA})$.[2] Kotlarski's aforementioned theorem was refined by Cieśliński [3] who proved that $\mathsf{CT}^-[\mathsf{PA}]+$ "T is closed under propositional proofs" and $\mathsf{CT}^-[\mathsf{PA}] + \mathsf{Ref}(\mathsf{PA})$ axiomatize the same theory. The main result of this paper, in turn, refines Cieśliński's result by demonstrating:

Theorem 1 $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}] + \mathsf{DC}$ *and* $\mathsf{CT}_0[\mathsf{PA}]$ *axiomatize the same first order theory.*

In the above theorem, $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}]$ is the weakening of $\mathsf{CT}^-[\mathsf{PA}]$ obtained by replacing the 'base theory' PA with its fragment consisting of Robinson's arithmetic Q, along with the scheme for Δ_0-induction and the totality of the exponential function; and DC (disjunctive correctness) is the statement asserting that a disjunction of

[1] CT[PA] dwarfs PA in arithmetical strength: By a classical theorem (discovered by a number of researchers, including Feferman and Takeuti, and explained in [9]) the arithmetical consequences of CT[PA] are the same as the arithmetical consequences of ACA, the subsystem of second order arithmetic obtained by adding the full scheme of induction over natural numbers (in the language of second order arithmetic) to the well-known subsystem ACA_0 of second order arithmetic.

[2] This result was first claimed by Kotlarski [11], but his proof outline of Ref(PA) within $\mathsf{CT}_0[\mathsf{PA}]$ was found to contain a serious gap in 2011 by Heck and Visser; this gap cast doubt over the veracity of Kotlarski's claim until the issue was resolved by Łełyk in his doctoral dissertation [15]. Łełyk's work was preceded by the discovery of an elegant proof of the nonconservativity of $\mathsf{CT}_0[\mathsf{PA}]$ over PA by Wcisło and Łełyk [16].

arithmetical sentences of arbitrary finite length is true (in the sense of T) iff one of the disjuncts is true. Coupled with Łełyk's aforementioned result [15], Theorem 1 shows that $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}] + \mathsf{DC}$, $\mathsf{CT}^-[\mathsf{PA}] + \mathsf{DC}$, and $\mathsf{CT}_0[\mathsf{PA}] + \mathsf{Ref}(\mathsf{PA})$ are axiomatizations of the same theory.

The plan of the paper is as follows: in Sect. 2 we review preliminary definitions and results, including more precise versions of those definitions and results mentioned in this introduction. In Sect. 3 we establish the veracity of the principle IC (Inductive Correctness, often referred to in the literature as "internal induction") within $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}] + \mathsf{DC}$. This is demonstrated by first establishing a new general form of Visser's theorem [20] on nonexistence of infinite descending chains of truth definitions with the help of (Löb's version of) Gödel's second incompleteness theorem instead of the Visser–Yablo paradox. In Sect. 4 we show that $\mathsf{CT}_0[\mathsf{PA}]$ is a subtheory of $\mathsf{CT}^-[\mathsf{PA}] + \mathsf{DC} + \mathsf{IC}$; thus Sects. 3 and 4 together constitute the proof of the hard direction of Theorem 1 since it is routine to verify that both DC and IC are theorems of $\mathsf{CT}_0[\mathsf{PA}]$. We close the paper with some open problems in Sect. 5.

Historical note The concept of disjunctive correctness first appeared in the work of Krajewski [13, p.133], who called it "$\vee$-completeness"; the current terminology was coined in a working paper of Enayat and Visser that was privately circulated in 2011, only a fragment [5] of which has been published so far. The working paper included the claim that $\mathsf{CT}^-[\mathsf{PA}] + \mathsf{DC}$ is conservative over PA, but the proof outline presented in the paper was found in 2013 to contain a significant gap by Cieśliński and his (then) doctoral students Łełyk and Wcisło. On the other hand, in 2012 Enayat found a proof of $\mathsf{CT}_0[\mathsf{PA}]$ within $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}] + \mathsf{DC} + \mathsf{IC}$; his proof was only privately circulated, and later was presented in the doctoral dissertation of Łełyk [15]. This proof forms the content of Sect. 4 of this paper. In light of these developments, and the well-known conservativity of $\mathsf{CT}^-[\mathsf{PA}] + \mathsf{IC}$ over PA (see Theorem 2.3), the question of conservativity of $\mathsf{CT}^-[\mathsf{PA}] + \mathsf{DC}$ over PA came to prominence amongst truth theory experts [4, p.226], and had been unsuccessfully attacked by a number of researchers since 2013, until Pakhomov established IC within $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}] + \mathsf{DC}$ as in Sect. 3 of this paper, which, coupled with Enayat's aforementioned result, yields Theorem 1 and exhibits the unexpected arithmetical strength of DC .

2 Preliminaries

Definition 2.1 (a) $\mathcal{L}_\mathsf{A}$ is the usual language of first order arithmetic $\{+, \cdot, \mathsf{S}(x), <, 0\}$. To simplify matters, we will assume that the logical constants of first order logic consist only of $\neg$ (negation), $\vee$ (disjunction), and $\exists$ (existential quantification); in particular $\forall$ (universal quantification) as well as $\wedge$ (conjunction) and other propositional connectives are treated here as derived notions.

(b) Given a language $\mathcal{L} \supseteq \mathcal{L}_\mathsf{A}$, an $\mathcal{L}$-formula φ is said to be a $\Delta_0(\mathcal{L})$-*formula* if all the quantifiers of φ are bounded by $\mathcal{L}$-terms, i.e., they are of the form $\exists x \leq t$, or of the form $\forall x \leq t$, where t is a term of $\mathcal{L}$ not involving x. Given a predicate $\mathsf{U}(x)$, $\mathcal{L}_{\mathsf{A}+\mathsf{U}}$ is the language $\mathcal{L}_\mathsf{A} \cup \{\mathsf{U}(x)\}$.

(c) Given a language $\mathcal{L} \supseteq \mathcal{L}_{\mathsf{A}}$, $\mathsf{I}\Delta_0(\mathcal{L})$ is the scheme of induction over natural numbers for $\Delta_0(\mathcal{L})$-formulae. *We shall omit the reference to $\mathcal{L}$ if $\mathcal{L} = \mathcal{L}_{\mathsf{A}}$, e.g., a Δ_0-formula is a $\Delta_0(\mathcal{L}_{\mathsf{A}})$-formula; and we shall use $\mathsf{I}\Delta_0(\mathsf{U})$ to abbreviate $\mathsf{I}\Delta_0(\mathcal{L}_{\mathsf{A}+\mathsf{U}})$.*

(d) $\mathsf{I}\Delta_0 + \mathsf{Exp}$ is the fragment of Peano arithmetic obtained by strengthening Robinson's arithmetic Q with $\mathsf{I}\Delta_0$ and with the sentence Exp that expresses the totality of the exponential function $y = 2^x$. It is well-known that Exp can be written as $\forall x \exists y\, \mathsf{Exp}(x, y)$, where $\mathsf{Exp}(x, y)$ is a Δ_0-predicate which, provably in $\mathsf{I}\Delta_0$, satisfies the familiar algebraic laws governing the graph of the exponential function, cf. [8, Sec. V3(c)].

(e) $\mathsf{Sent}_{\mathsf{A}}(x)$ is the $\mathcal{L}_{\mathsf{A}}$-formula that expresses "x is the Gödel-number of a formula of $\mathcal{L}_{\mathsf{A}}$ with no free variables", and $\mathsf{Form}^n_{\mathsf{A}}(x)$ is the $\mathcal{L}_{\mathsf{A}}$-formula that expresses "x is the Gödel-number of a formula of $\mathcal{L}_{\mathsf{A}}$ with precisely n free variables". We use $\mathsf{Sent}_{\mathsf{A}}$ and $\mathsf{Form}^n_{\mathsf{A}}$ to refer to the corresponding definable classes of Gödel-numbers of $\mathcal{L}_{\mathsf{A}}$-formulae.

(f) Given a (base) theory B whose language is $\mathcal{L}_{\mathsf{A}}$ and which extends $\mathsf{I}\Delta_0 + \mathsf{Exp}$, $\mathsf{CT}^-[\mathsf{B}]$ is the theory obtained by strengthening B by adding the sentences tarski_0 through tarski_4 described below, where we use the following conventions: τ_1 and τ_2 vary over Gödel-numbers of closed $\mathcal{L}_{\mathsf{A}}$-terms, τ_i° is the value of the term coded by τ_i, φ and ψ range over Gödel-numbers of $\mathcal{L}_{\mathsf{A}}$-sentences, v ranges over variables, $\gamma(v)$ ranges over $\mathsf{Form}^1_{\mathsf{A}}$, and $\underline{x}$ is the numeral representing the value of x.

$$\begin{aligned}
&\mathsf{tarski}_0 := \forall x\, (\mathsf{T}(x) \to \mathsf{Sent}_{\mathsf{A}}(x)).\\
&\mathsf{tarski}_1 := \forall \tau_1, \tau_1\ \big(\mathsf{T}(\tau_1 = \tau_2) \leftrightarrow \tau_1^\circ = \tau_2^\circ\big).\\
&\mathsf{tarski}_2 := \forall \varphi((\mathsf{T}(\neg\varphi) \leftrightarrow \neg\mathsf{T}(\varphi)).\\
&\mathsf{tarski}_3 := \forall \varphi, \psi\ \ (\mathsf{T}(\varphi \vee \psi) \leftrightarrow (\mathsf{T}(\varphi) \vee \mathsf{T}(\psi))).\\
&\mathsf{tarski}_4 := \forall v\, \forall \gamma(v)\ (\mathsf{T}\,(\exists v\, \gamma(v)) \leftrightarrow \exists x\, \mathsf{T}(\gamma(\underline{x})).
\end{aligned}$$

(g) $\mathsf{CT}_0[\mathsf{B}] := \mathsf{CT}^-[\,\mathsf{B}] \cup \mathsf{I}\Delta_0(\mathsf{T})$.

(h) DC (disjunctive correctness) is the $\mathcal{L}_{\mathsf{A}+\mathsf{T}}$-sentence asserting that T commutes with disjunctions of arbitrary length, i.e., DC asserts that for all numbers s and for all sequences $\langle \varphi_i : i < s\rangle$ from $\mathsf{Sent}_{\mathsf{A}}$, the following equivalence holds:

$$\mathsf{T}\left(\bigvee_{i<s} \varphi_i\right) \leftrightarrow \exists i < s\ \mathsf{T}(\varphi_i),$$

where for definiteness $\bigvee_{i<s} \varphi_i$ is defined[3] by recursion: $\bigvee_{i<0} \varphi_i := \varphi_0$ and $\bigvee_{i<t+1} \varphi_i := \big(\bigvee_{i<t} \varphi_i\big) \vee \varphi_t$.

(i) We will employ the abbreviation $\bigwedge_{i<s} \varphi_i$ for $\neg \bigvee_{i<s} \neg\varphi_i$, and CC (conjunctive correctness) for the $\mathcal{L}_{\mathsf{A}+\mathsf{T}}$-sentence

$$\mathsf{T}\left(\bigwedge_{i<s} \varphi_i\right) \leftrightarrow \forall i < s\ \mathsf{T}(\varphi_i).$$

[3] One can also formulate disjunctive correctness in a stronger way by asserting that all disjunctions (and not just the ones that are grouped to the left) are well-behaved with respect to T. But the current frugal form, as shown by Theorem 1, ends up implying the seemingly stronger form since it is easy to show in CT_0 that the two forms are equivalent.

- Note that the commutativity of T with negation implies that DC and CC are equivalent.

(j) IC (inductive correctness[4]) is the $\mathcal{L}_{\mathsf{A}+\mathsf{T}}$-sentence asserting that each $\mathcal{L}_{\mathsf{A}}$-instance of induction over natural numbers is true, i.e., IC asserts that for all unary $\mathcal{L}_{\mathsf{A}}$-formulae $\psi = \psi(x)$, $\mathsf{T}(\ulcorner \mathsf{Ind}_{\psi} \urcorner)$ holds, where Ind_{ψ} is the following $\mathcal{L}_{\mathsf{A}}$-sentence that asserts that ψ is inductive:

$$\psi(\underline{0}) \rightarrow (\forall x\, (\psi(x) \rightarrow \psi(x+1)) \rightarrow \forall x\, \psi(x))\,.$$

The $\mathsf{B} = \mathsf{PA}$ case of Theorem 2.2 below, and its elaboration Theorem 2.3, were first established in the work of Krajewski et al. [12] for $\mathsf{B} = \mathsf{PA}$, where PA is formulated in a relational language, and 'domain constants' play the role of numerals. As mentioned in the introduction to this paper, their result was couched in model theoretic terms involving the notions of recursive saturation and satisfaction classes, but it is well-known that their formulation is equivalent to appropriately formulated conservativity assertions (the key ingredients of this equivalence are the following facts: Every consistent theory in a countable language has a recursively saturated model, and countable recursively saturated models are resplendent). Later Kaye [10] developed the theory of satisfaction classes over models of PA in languages incorporating function symbols; his work was extended by Engström [6] to truth classes over models of PA in functional languages.[5] More recently, newer and more informative proofs of Theorems 2.2 and 2.3 have been found in the joint work of Visser and Enayat [5] (with base theories that support a modicum of coding, and which are formulated in a relational language), and by Leigh [14] (for functional base theories that support a modicum of coding). As verified by Cieśliński [4, Ch. 7], the Visser-Enayat model theoretic methodology can be extended so as to accommodate functional languages.

Theorem 2.2 $\mathsf{CT}^{-}[\mathsf{B}]$ *is conservative over* B *for every arithmetical base theory* B *extending* $\mathsf{I}\Delta_0 + \mathsf{Exp}$.

Theorem 2.3 $\mathsf{CT}^{-}[\mathsf{PA}] + \mathsf{IC}$ *is conservative over* PA.

The direction (a) $\Rightarrow$ (b) of Theorem 2.4 below is due to Kotlarski [11]; the other direction is due to Łełyk [15].

Theorem 2.4 (Kotlarski–Łełyk) *The following theories are deductively equivalent:*

(a) $\mathsf{CT}^{-}[\mathsf{PA}] + \mathsf{Ref}(\mathsf{PA})$.
(b) $\mathsf{CT}_0[\mathsf{PA}]$.

The direction (a) $\Rightarrow$ (b) of Theorem 2.5 below is due to Cieśliński [3], who refined Kotlarski's proof of the direction (a) $\Rightarrow$ (b) of Theorem 2.4; the other direction involves a routine induction.

Theorem 2.5 (Cieśliński) *The following theories are deductively equivalent:*

(a) $\mathsf{CT}^{-}[\mathsf{PA}] +$ "T *is closed under propositional proofs*".
(b) $\mathsf{CT}_0[\mathsf{PA}]$.

[4] This condition has been referred to as Int (internal induction) in the literature.

[5] The subtle distinction between satisfaction classes and truth classes, and their close relationship, is explained in [5] and [4, Ch. 7].

3 Disjunctive correctness implies inductive correctness

Definition 3.1 ITB (iterated truth biconditionals) is a theory formulated in two-sorted first order logic. The first sort $x, y, z \ldots$ of ITB is for the 'natural numbers'. The second sort $\alpha, \beta, \gamma, \ldots$ is for the *indices* of truth definition. The language $\mathcal{L}_{\mathsf{ITB}}$ of ITB is obtained by augmenting the language $\mathcal{L}_{\mathsf{A}}$ of arithmetic with two binary predicates: $\alpha \prec \beta$ and $\mathsf{T}(\alpha, x)$, but we shall write $\mathsf{T}(\alpha, x)$ as $\mathsf{T}_\alpha(x)$ to display the indexicality of α. The axioms of ITB come in three groups. The first group consists of the axioms of Q (Robinson arithmetic); the second group consists of a single axiom asserting that $\prec$ is a transitive relation; and the third group consists of the following biconditionals B_φ:

$$\mathsf{B}_\varphi := \forall \alpha \left(\mathsf{T}_\alpha(\ulcorner \varphi \urcorner) \leftrightarrow \varphi^{\prec \alpha} \right),$$

where φ ranges over all $\mathcal{L}_{\mathsf{ITB}}$ -sentences, and for each index variable α, $\varphi^{\prec \alpha}$ denotes the relativization of φ to the cone of indices below α, i.e. the formula obtained by replacing all the quantifiers of the form $\forall \beta$ $(\exists \beta)$ with $\forall \beta \prec \alpha$ $(\exists \beta \prec \alpha)$, and if there is a bounded instance of α we make the appropriate renaming. Clearly $\varphi^{\prec \alpha} = \varphi$ if φ is a purely arithmetical formula.

- Note that we take the theory ITB over the variant of many-sorted logic that allows domains of some sorts to be empty.
- Although we haven't required $\prec$ to be irreflexive, we treat it as an irreflexive relation; in particular we define a minimal element α to be an element such that $\forall \beta \neg(\beta \prec \alpha)$. The reason is that Theorem 3.2 below implies that ITB proves the irreflexivity of $\prec$. Alternatively one could show that ITB proves irreflexivity of $\prec$ by observing that the existence of a model of ITB with a reflexive point contradicts Tarksi's undefinability of truth theorem.
- We will use the following convention to lighten the notation: The notation $\ulcorner \varphi \urcorner$ for the Gödel number of a formula φ will be generally used, but the corner-notation will be omitted when φ appears inside of a truth predicate T, or inside an indexed version of T.

The proof of the following theorem was inspired by the recent James Walsh proof [18] of nonexistence of infinite recursive provably descending chains of sentences with respect to $<_{\mathsf{Con}}$-order. We note that Theorem 3.2 is similar to a result by Flumini and Sato [7], which states that the second order principle asserting the existence of iteration of Π^0_1-comprehension over a preorder $\prec$ implies that $\prec$ is well-founded.

Theorem 3.2 $\mathsf{ITB} + \exists \alpha(\alpha = \alpha)$ *proves the existence of a* $\prec$*-minimal element. Equivalently, the following theory* DTB *(descending truth biconditionals) is inconsistent:*

$$\mathsf{DTB} := \mathsf{ITB} + \forall \alpha \exists \beta(\beta \prec \alpha) + \exists \alpha(\alpha = \alpha).$$

Proof We prove the inconsistency of DTB by Löb's version of Gödel's second incompleteness theorem[6]: We exhibit a formula $\theta(x)$ that satisfies the HBL (Hilbert–Bernays–Löb) conditions for a provability predicate over the theory DTB. This enables

[6] Löb's paper [17], in which the venerable 'Löb's Theorem' was proved, is responsible for the now common standard textbook framework for the presentation of 'abstract' form of Gödel's second incompleteness

us to justify the inconsistency of DTB by showing that DTB proves the "consistency" sentence $\neg\theta(0=1)$.

Consider the formula $\theta(x)$:

$$\theta(x) := \forall\alpha(\mathsf{T}_\alpha(x)).$$

We will verify that $\theta(x)$ satisfies the HBL conditions listed below provably in DTB; in what follows φ and ψ range over all sentences of the language of DTB:

HBL-1. $\mathsf{DTB} \vdash \varphi \Longrightarrow \mathsf{DTB} \vdash \theta(\ulcorner\varphi\urcorner)$.
HBL-2. $\mathsf{DTB} \vdash \theta(\ulcorner\varphi \to \psi\urcorner) \to (\theta(\ulcorner\varphi\urcorner) \to \theta(\ulcorner\psi\urcorner))$.
HBL-3. $\mathsf{DTB} \vdash \theta(\ulcorner\varphi\urcorner) \to \theta(\ulcorner\theta(\ulcorner\varphi\urcorner)\urcorner)$.

Since we have biconditionals, in order to prove HBL-1 it is enough to show that for each sentence φ, if $\mathsf{DTB} \vdash \varphi$ then $\mathsf{DTB} \vdash \forall\alpha\ \varphi^{\prec\alpha}$. The latter is the case since for any model $\mathcal{M}$ of DTB and index a in $\mathcal{M}$, the theory DTB holds in the model $\mathcal{M}^{\prec a}$ that is the restriction of $\mathcal{M}$ to all indices $\prec a$.

For a given φ and ψ, HBL-2 follows directly from the biconditional axioms $\mathsf{B}_{\varphi\to\psi}$, B_φ, and B_ψ of ITB.

Finally, HBL-3 holds since:

$$\mathsf{ITB} \vdash \theta(\ulcorner\theta(\ulcorner\varphi\urcorner)\urcorner) \longleftrightarrow \left(\forall\alpha\forall\beta((\beta \prec \alpha) \to \mathsf{T}_\beta(\varphi))\right).$$

On the other hand, the formula $\forall\alpha\ \neg\mathsf{T}_\alpha(\ulcorner\underline{0} = \underline{1}\urcorner)$ is provable in ITB, hence the formula $\neg\theta(\ulcorner\underline{0} = \underline{1}\urcorner)$ is provable in ITB, and therefore in $\mathsf{ITB}+\exists\alpha(\alpha=\alpha)$. So by Löb's version of Gödel's second incompleteness theorem, DTB is inconsistent. □

Lemma 3.3 $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}] + \mathsf{DC}$ *proves* IC.

Proof By Theorem 3.2 we can fix an inconsistent finite subtheory DTB^- of DTB. Suppose DTB^- contains only biconditionals for the formulae $\varphi_0, \ldots, \varphi_{k-1}$. We will use ITB^- to denote the subtheory of ITB whose only biconditional axioms are $\left\{\mathsf{B}_{\varphi_i} : i < k\right\}$.

For the rest of the proof we will reason in $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}] + \mathsf{DC}$. In order to prove IC we assume for a contradiction that some arithmetical $\psi(x)$ is not inductive in the sense of T, i.e., we have:

$$\neg\mathsf{T}\left(\psi(0) \to (\forall x\ (\psi(x) \to \psi(x+1)) \to \forall x\ \psi(x))\right).$$

Within $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}] + \mathsf{DC}$ we use induction on n to define translations ι_n from the language of ITB to the language of first-order arithmetic such that from the point of view of T all of them will be interpretations of ITB^-, i.e., we will have $\mathsf{T}(\iota_n(\varphi))$, for all axioms φ of ITB^-. We will arrive at a contradiction by showing that $\neg\mathsf{T}(\psi(\underline{n}))$

Footnote 6 continued
theorem: If T is a consistent theory extending Q that supports a unary predicate $\theta(x)$ satisfying conditions HBL-2, HBL-2, and HBL-3, then T doesn't prove $\theta(\ulcorner 0 = 1\urcorner)$, i.e., the consistency sentence corresponding to θ (the intended meaning of $\theta(x)$ is "the sentence with Gödel number x is provable in T"). See, e.g., [1, Ch. 18], for the presentation of such a general form of Gödel's second incompleteness theorem.

implies that ι_n is an interpretation of DTB^-; thus it will be necessary to consider n that are non-standard from external point of view.

Note that each translation ι_n consists of finitely many formulae, giving the interpretation of the domains and all symbols of the signature of ITB, and thus could be easily represented by a number. The interpretation of arithmetic in each ι_n is the identity interpretation, but the domain of indices of truth definition ι_n is given by the following formula $D^{(n)}(x)$:

$$x < \underline{n} \wedge \neg\psi(x).$$

For all n the relation $\prec$ is interpreted by $<$. The formula $\mathsf{T}_\alpha(x)$ is interpreted by the formula $\mathsf{IT}^{(n)}(y, x)$, where y corresponds to α, and x corresponds to itself:

$$\bigwedge_{i<k}(x = \varphi_i \to \bigwedge_{m<n}((y = \underline{m} \wedge \neg\psi(\underline{m})) \to \iota_m(\varphi_i))),$$

where $\iota_m(\varphi_i)$ is the ι_m-translation of the sentence φ_i. It is easy to see that this definition could be carried out in $\mathsf{I}\Delta_0 + \mathsf{Exp}$.

Let us now prove that the translations given by ι_n are indeed the desired interpretations inside T, i.e., we need to prove that for all n and axioms A of ITB^- we have $\mathsf{T}(\iota_n(\mathsf{A}))$. Clearly it is the case for all the axioms of Q and the axioms of partial order for $\prec$. Now let us show that for any $s < k$:

$$\mathsf{T}(\iota_n(\forall\alpha(\mathsf{T}_\alpha(\varphi_s) \leftrightarrow \varphi_s^{\prec\alpha}))). \qquad (\blacktriangle)$$

By compositional axioms, we just need to show that for all u such that $u < n$ and $\mathsf{T}(\neg\psi(\underline{u}))$ we have:

$$\mathsf{T}\left(\bigwedge_{i<k}\left(\varphi_s = \varphi_i \to \bigwedge_{m<n}\left(\left(\underline{u} = \underline{m} \wedge \neg\psi(\underline{m})\right) \to \iota_m(\varphi_i)\right)\right)\right) \leftrightarrow \mathsf{T}\left(\iota_n\left(\varphi_s^{\prec\underline{u}}\right)\right).$$

Now by compositional axioms and DC (in the form of CC, as explained in part (i) of Definition 2.1) our task can be reduced to proving the equivalence:

$$\mathsf{T}\left(\iota_u\left(\varphi_s\right)\right) \leftrightarrow \mathsf{T}\left(\iota_n\left(\varphi_s^{\prec\underline{u}}\right)\right).$$

In order to prove this we will show by induction on subformulae θ of φ_s that for the universal closure $\overline{\theta}$ of θ:

$$\mathsf{T}\left(\iota_u(\overline{\theta})\right) \leftrightarrow \mathsf{T}\left(\iota_n(\overline{\theta}^{\prec\underline{u}})\right).$$

Note that since φ_s is a fixed formula with finitely many subformulae, actually this external induction will be formalizable in $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}] + \mathsf{DC}$ despite the fact that it lacks the induction axiom for the appropriate class of formulae. The only non-trivial case here is the case when θ is $\mathsf{T}_\alpha(x)$:

$$\mathsf{T}\left(\iota_u\left(\forall\alpha\forall x\mathsf{T}_\alpha(x)\right)\right) \leftrightarrow \mathsf{T}(\iota_n(\forall\alpha \prec \underline{u}\ \forall x(\mathsf{T}_\alpha(x)))).$$

Hence we just need to show that for all $p < u$ such that $\mathsf{T}(\neg\psi(\underline{p}))$, and for all o, the following pair of formulae (whose only formal difference is in the bound for indices of the second conjunction) are equivalent:

$$\mathsf{T}\left(\bigwedge_{i<k}\left(\underline{o}=\varphi_i \rightarrow \bigwedge_{m<u}\left(\left(\underline{p}=\underline{m}\wedge\neg\psi(\underline{m})\right)\rightarrow \iota_m(\varphi_i)\right)\right)\right),$$

$$\mathsf{T}\left(\bigwedge_{i<k}\left(\underline{o}=\varphi_i \rightarrow \bigwedge_{m<n}\left(\left(\underline{p}=\underline{m}\wedge\neg\psi(\underline{m})\right)\rightarrow \iota_m(\varphi_i)\right)\right)\right).$$

But since $p < u < n$, we trivially use DC (in the form of CC) to show that the formulae are indeed equivalent. Thus we conclude that (▲) holds.

Choose some n such that $\mathsf{T}(\neg\psi(\underline{n}))$; this is possible since we assumed that induction fails for $\psi(x)$ in the sense of T. It is easy to see that ι_n actually is an interpretation of DTB^- inside T. We externally fix some proof of inconsistency from axioms of DTB^- and follow it inside T to derive a contradiction, thereby completing the proof of IC . □

Corollary 3.4 $\mathsf{CT}^-[\,\mathsf{I}\Delta_0+\mathsf{Exp}]+\mathsf{DC}$ *proves* PA*, and therefore* $\mathsf{CT}^-[\mathsf{I}\Delta_0+\mathsf{Exp}]+\mathsf{DC}$ *and* $\mathsf{CT}^-[\mathsf{PA}]+\mathsf{DC}$ *axiomatize the same theory.*

Proof This is an immediate consequence of Lemma 3.3, and the provability of Tarski bi-conditionals in $\mathsf{CT}^-[\mathsf{I}\Delta_0+\mathsf{Exp}]$. □

Remark 3.5 Note that Theorem 3.2 could be regarded as a strengthening of Tarski's undefinability of truth theorem: Tarski's theorem essentially states that there could be no hierarchy of truth definitions whose set of indices contains a reflexive point. To the best of the authors' knowledge there is no known proof of Tarski's theorem that avoids the use of any kind of diagonalization constructions.[7] In a preprint of this paper we raised a question of whether Lemma 3.3, which we proved using Theorem 3.2, could be proved more directly without the use of diagonalization. Later we noticed that it is possible to replace the use of Theorem 3.2 in the proof of Lemma 3.3 with a result by Flumini and Sato [7, Thm. 1], which has a rather simple proof that in our opinion could be regarded as diagonalization-free. Since this other proof is of methodological interest, we will provide its general outline.

The results of Flumini and Sato in [7] are fairly general and are applicable to various second-order systems. To keep our presentation compact we will just formulate a direct corollary of [7, Thm. 1] that will be relevant to us. The only axioms of our base system of second-order arithmetic B will be those of $\mathsf{I}\Delta_0+\mathsf{Exp}$. We denote as Ind the usual second-order induction principle:

$$\forall X\ (0\in X\wedge\forall x\ (x\in X\rightarrow x+1\in X)\rightarrow\forall x\ x\in X).$$

For a set X and number a we denote as $(X)_a$ the set $\{b \mid \langle a,b\rangle\in X\}$ and for binary relation $\prec$ we denote as $X^{\prec a}$ the set $\{\langle a',b\rangle \mid \langle a',b\rangle\in X \text{ and } a'\prec a\}$. For a second

[7] We refer to the introduction of Visser's paper [21] for a discussion of the role of diagonalization in the proofs of Gödel's second incompleteness theorem and Tarski's truth undefinability theorem.

order formula $\varphi(X, y)$ and binary relation $\prec$ we could naturally write the formula $\mathsf{Hier}(\varphi, \prec, H)$ that expresses that H is a hierarchy along $\prec$ produced by φ, i.e., H is such that $\forall x\ ((H)_x = \{y \mid \varphi(H^{\prec x}, y)\})$. The usual arithmetic transfinite recursion principle states that hierarchies exist for any well-ordering $\prec$ and formula φ without second-order quantifiers. The corollary of [7, Thm. 1] that will be relevant to us is that over B there is an arithmetic (moreover Π^0_1) formula $\varphi(X, y)$ (with additional variables) such that the axiom Ind is implied by the existence of φ-hierarchies along arbitrarily large proper initial segments of natural numbers. Formally this principle is $\forall z \exists H\ \mathsf{Hier}(\varphi, <\restriction z, H)$, where $<\restriction z$ is the restriction of the usual order $<$ on naturals to the numbers below z.

Now consider the ω-interpretation of the language of second-order arithmetic in $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}]$, where the range of sets consists of all the sets $\{n \mid \mathsf{T}(\varphi(\underline{n}))\}$. It is easy to see that IC is equivalent to the validity of Ind in this interpretation.[8] So in view of the above in order to show that DC implies Ind it will be enough to prove that DC implies that the universal closure of $\forall z \exists H\ \mathsf{Hier}(\varphi, <\restriction z, H)$ holds in this interpretation. To achieve the latter we could just directly construct the formula defining the relevant hierarchy and then use DC to verify that it indeed has the desired property.

4 Disjunctive correctness + inductive correctness implies $\mathbf{\Delta_0(T)}$-induction

In this section we shall prove that $\mathsf{I}\Delta_0(\mathsf{T})$ is provable in $\mathsf{CT}^-[\mathsf{PA}] + \mathsf{DC} + \mathsf{IC}$, which, coupled with Lemma 3.3 completes the proof of the nontrivial direction of Theorem 1. We begin with a key definition:

Definition 4.1 In what follows $\in_{\mathsf{Ack}}$ is "Ackermann's epsilon", i.e., $x \in_{\mathsf{Ack}} y$ is the arithmetical formula that expresses "the x-th bit of the binary expansion of y is a 1".

(a) For a unary predicate $\mathsf{U}(x)$, the $\mathcal{L}_{\mathsf{A}+\mathsf{U}}$ sentence PC_{U} (read as "U is piece-wise coded") is the following sentence:

$$\forall u\ \exists y\ \forall x\ [(\mathsf{U}(x) \wedge x < u) \leftrightarrow x \in_{\mathsf{Ack}} y]\,.$$

(b) More generally, given an n-ary $\mathcal{L}_{\mathsf{A}+\mathsf{U}}$-formula $\varphi(\mathsf{U}, x_0, \ldots, x_{n-1})$, PC_φ is the following $\mathcal{L}_{\mathsf{A}+\mathsf{U}}$-sentence:

$$\forall u\ \exists y\ \forall x_0, \ldots \forall x_{n-1}\ \big[(\varphi(\mathsf{U}, x_0, \ldots, x_{n-1}) \wedge (x_0 < u \wedge \cdots \wedge x_{n-1} < u) \leftrightarrow \langle x_i : i < n\rangle \in_{\mathsf{Ack}} y\big],$$

where $\langle x_i : i < n\rangle$ is a canonical code for the ordered n-tuple $(x_0, \ldots, x_{n-1})$.

The following lemma shows that over $\mathsf{I}\Delta_0 + \mathsf{Exp}$ the scheme $\mathsf{I}\Delta_0(\mathsf{U})$ is equivalent to the single sentence 'U is piecewise coded'. The lemma is folklore; we present the proof for the sake of completeness.

[8] See Remark 4.3.1 for a sharper formulation of this equivalence.

Lemma 4.2 *The following are equivalent over* $\mathsf{I}\Delta_0 + \mathsf{Exp}$*:*

(a) $\mathsf{I}\Delta_0(\mathsf{U})$.
(b) PC_{U}.

Proof We will reason in $\mathsf{I}\Delta_0 + \mathsf{Exp}$. Recall that both $x = 2^y$ and $x \in_{\mathsf{Ack}} y$ have Δ_0-definitions within $\mathsf{I}\Delta_0$ [8, Ch. V].

(**a** $\longrightarrow$ **b**): Assume $\mathsf{I}\Delta_0(\mathsf{U})$. Given u, let w be the Ackermann-code for the set of predecessors of u (i.e., $\forall x\,(x < u \leftrightarrow x \in_{\mathsf{Ack}} w)$). Clearly $w = \sum_{i<u} 2^i = 2^u - 1$, and w is an upper bound for any w' that codes a subset of the predecessors of u. Let $\delta(u, w)$ be the Δ_0-formula below:

$$(w = 2^u) \rightarrow [\exists y < w\ \forall x < u\ \ ((\mathsf{U}(x) \wedge x < u) \leftrightarrow x \in_{\mathsf{Ack}} y)]\,.$$

A simple induction on u (where w is treated as parameter) using $\mathsf{I}\Delta_0(\mathsf{U})$ shows that $\forall u\ \forall w\ \delta(u, w)$ holds, which completes the proof that PC_{U} holds.
(**b** $\rightarrow$ **a**): A straightforward induction on the complexity of $\Delta_0(\mathsf{U})$-formulae shows that:

$(*)$ If U is piecewise coded and $\delta(\mathsf{U}, x_0, \ldots, x_{n-1})$ is a $\Delta_0(\mathsf{U})$-formula, then PC_δ holds.

The base case of the induction is clearly equivalent to the assumption that U is piecewise coded. What allows the inductive steps to be smoothly carried out is that, provably in $\mathrm{I}1_0 + \mathrm{Exp}$, $\in_{\mathsf{Ack}}$ obeys many familiar axioms of set theory, as verified in [8, Ch. I, Thm. 1.39]. As an example, in the existential case of the induction, we suppose that $\delta(\mathsf{U}, x_0, \ldots, x_{n-1})$ is a $\Delta_0(\mathsf{U})$-formula such that PC_δ holds, and then establish $\mathsf{PC}_{\delta'}$, where $\delta' = \exists x_0 < t(x_0, \ldots, x_{n-1})\ \delta(\mathsf{U}, x_0, \ldots, x_{n-1})$ for some term t. To do so, let us fix any number u and demonstrate that there is an $\in_{\mathsf{Ack}}$-set s', where:

$$s' = \{\langle x_i : 0 < i < n\rangle \mid x_1, \ldots, x_{n-1} < u \ \text{ and } \ \delta'(U, x_1, \ldots, x_{n-1})\}.$$

Let v be a number such that the value $t(x_0, \ldots, x_{n-1}) \leq v$ for $x_0, \ldots, x_{n-1} < u$. By PC_δ we have the following $\in_{\mathsf{Ack}}$-set s:

$$s = \{\langle x_i : i < n\rangle \mid x_0, \ldots, x_{n-1} < v \ \text{ and } \ \delta(U, x_0, \ldots, x_{n-1})\}.$$

Using Δ_0-Separation we construct the set s' as:

$$\{\langle x_i : 0 < i < n\rangle \mid x_1, \ldots, x_{n-1} < u \text{ and } \exists x_0 < t(x_0, \ldots, x_{n-1})\ \ \langle x_i : i < n\rangle \in_{\mathsf{Ack}} s)\}.$$

With $(*)$ at our disposal, we could trivially deduce the least number principle for $\Delta_0(\mathsf{U})$-formula, which is of course equivalent to $\mathsf{I}\Delta_0(\mathsf{U})$. □

- In the next lemma and its proof, $\mathsf{Code}(c, \varphi, u)$ denotes the ternary $\mathcal{L}_{\mathsf{A}+\mathsf{T}}$-formula $\forall x \left[\left(x < u \wedge \mathsf{T}\left(\varphi(\underline{x})\right)\right) \leftrightarrow x \in_{\mathsf{Ack}} c\right]$,
and $\mathsf{PC}(\varphi)$ denotes the $\mathcal{L}_{\mathsf{A}+\mathsf{T}}$-formula $\forall u\ \exists c\ \mathsf{Code}(c, \varphi, u)$.

Lemma 4.3 $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}] + \mathsf{IC}$ *proves* $\forall\varphi(\mathsf{Form}^1_{\mathsf{A}}(\varphi) \to \mathsf{PC}(\varphi))$.

Proof We reason in $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}] + \mathsf{IC}$. Given $\varphi(x)$ in $\mathsf{Form}^1_{\mathsf{A}}$, we need to show:

(1) $\forall u\, \exists c\, \mathsf{Code}(c, \varphi, u)$.
    By the compositional properties of T, (1) is equivalent to:
(2) $\mathsf{T}(\forall u\, \psi(u))$, where $\psi(u) := \exists c(\forall x < u\ \varphi(x) \leftrightarrow x \in_{\mathsf{Ack}} c)$.

On the other hand, $\forall u\ \psi(u)$ is the conclusion of the formula Ind_ψ (asserting the inductive property of ψ) given by IC. So (2) follows directly from IC and the easily verified facts $\mathsf{T}(\psi(\underline{0}))$ and $\mathsf{T}(\forall u\ \ (\psi(u) \to \psi(u+1)))$. □

Remark 4.3.1 Lemma 4.3 can be readily strengthened to a more general result whose proof we leave to the reader: $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}] + \mathsf{IC}$ verifies ACA_0 for the ω-interpretation of the language of second-order arithmetic, where the range of sets consists of all the sets $\{n \mid \mathsf{T}(\varphi(\underline{n}))\}$, where φ is any first-order formula. Moreover, it is a theorem of $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}]$ that IC is equivalent to the veracity of ACA_0 within this interpretation.

Lemma 4.4 $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}] + \mathsf{DC} + \mathsf{IC} \vdash \mathsf{I}\Delta_0(\mathsf{T})$.

Proof Reason in $\mathsf{CT}^-[\mathsf{I}\Delta_0 + \mathsf{Exp}] + \mathsf{DC} + \mathsf{IC}$. By Lemma 4.2, it suffices to show that T is piecewise coded. Let $\langle \varphi_i : i < u\rangle$ be the sequence of arithmetical sentences such that φ_i is the sentence with Gödel-number i if there is such a sentence, and otherwise φ_i is the sentence $\underline{0} = \underline{1}$. We wish to show that $\{i < u : \mathsf{T}(\varphi_i)\}$ is coded. Towards this goal, consider the unary formula $\theta(x) \in \mathsf{Form}_{\mathsf{A}}$ given by:

$$\theta(x) := \bigvee_{i<u} \left((x = \underline{i}) \wedge \varphi_i\right).$$

Claim $(*)$ $\forall i < u\ \left[\mathsf{T}(\varphi_i) \leftrightarrow \mathsf{T}(\theta(\underline{i}))\right]$.
$(\rightarrow)$ Suppose $\mathsf{T}(\varphi_i)$ for some $i < u$. Then $\mathsf{T}\left((\underline{i} = \underline{i}) \wedge \varphi_i\right)$, and hence by DC we have $\mathsf{T}(\theta(\underline{i}))$.
$(\leftarrow)$ Suppose $\mathsf{T}(\theta(\underline{i}))$ for some $i < u$. Then by DC, there is some $j < u$ such that $\mathsf{T}\left((\underline{i} = \underline{j}) \wedge \varphi_j\right)$. So $\mathsf{T}(\varphi_i)$ holds since T commutes with conjunction and $\mathsf{T}(\underline{i} = \underline{j})$ holds iff $i = j$.
By coupling Claim $(*)$ together with Lemma 4.3, we can conclude that $\left\{i < u : \mathsf{T}\left(\theta(\underline{i})\right)\right\}$ is coded. □

5 Closing remarks and open questions

Question 5.1 *Is the generalization of Theorem 1 in which* CT^- *is weakened to* CS^- *(where* S *stands for satisfaction) true?*

- The notion $\mathsf{CS}^-[\mathsf{B}]$ is defined in [5] for base theories B formulated in relational languages, using the notation B^{FS} (FS for "full satisfaction"); and in [4, Ch. 7] for functional languages. We expect that an examination of the proofs in Sects. 3 and 4 would show that this question has a positive answer.

Question 5.2 *Is* IC *provable in* $\mathsf{CT}^-[\mathsf{S}^1_2] + \mathsf{DC}$?

- In the above, S^1_2 is Buss's well-known arithmetical theory whose provable recursive functions are precisely the functions computable in polynomial time, as in [2]. For the above question to make sense, part (f) of Definition 1.1 should be adjusted so as to accommodate the fact that the language of S^1_2 extends $\mathcal{L}_\mathsf{A}$. In the proof of Lemma 3.3, most likely it is possible to use some tricks with effective formulae (see [19, Section 3]) in order to modify the definition of ι_n in such a way that their sizes will be polynomial. But in order for the construction to work we will also need to ensure that DC is still enough to show that ι_n are indeed interpretations inside the truth predicate.

Question 5.3 *Let* $\mathsf{DC}_\mathsf{Elim}$ *be the 'half' of* DC *that asserts: if a finite disjunction is true, then at least one of the disjuncts is also true. Is* $\mathsf{CT}^-[\mathsf{PA} + \mathsf{DC}_\mathsf{Elim}]$ *conservative over* PA*?*

- DC can be written as the conjunction of two implications $\mathsf{DC}_\mathsf{Elim}$ and $\mathsf{DC}_\mathsf{Intro}$, where $\mathsf{DC}_\mathsf{Intro}$ is the converse of $\mathsf{DC}_\mathsf{Elim}$ (i.e., $\mathsf{DC}_\mathsf{Intro}$ asserts: a disjunction is true, whenever at least one of its disjuncts is true). Recent joint of work of Wcisło, Łełyk and Enayat (to appear) shows that $\mathsf{DC}_\mathsf{Intro}$ can be conservatively added to $\mathsf{CT}^-[\mathsf{PA}] + \mathsf{IC} + \left\{\forall x \left(\mathsf{True}_{\Sigma_n}(x) \to \mathsf{T}(x)\right) : n < \omega\right\}$.

References

1. Boolos, G.S., Burgess, J.P., Jeffrey, R.C.: Computability and Logic, 5th edn. Cambridge University Press, Cambridge (2007)
2. Buss, S.: Proof theory of arithmetic. In: Buss, S. (ed.) Handbook of Proof Theory, pp. 79–147. Elsevier, Amsterdam (1998)
3. Cieśliński, C.: Deflationary truth and pathologies. J. Philos. Logic **39**, 325–337 (2010)
4. Cieśliński, C.: The Epistemic Lightness of Truth. Deflationism and its Logic. Cambridge University Press, Cambridge (2017)
5. Enayat, A., Visser, A.: New constructions of full satisfaction classes. In: Achourioti, T., Galinon, H., Fujimoto, K., Martínez-Fernández, J. (eds.) Unifying the Philosophy of Truth, pp. 321–325. Springer, Berlin (2015)
6. Engström, F.: Satisfaction classes in nonstandard models of first-order arithmetic (2002). arXiv:math/0209408
7. Flumini, D., Sato, K.: From hierarchies to well-foundedness. Arch. Math. Log. **54**, 855–863 (2014)
8. Hájek, P., Pudlák, P.: Metamathematics of First-Order Arithmetic. Springer, Berlin (1993)
9. Halbach, V.: Axiomatic Theories of Truth, 2nd edn. Cambridge University Press, Cambridge (2015)
10. Kaye, R.: Models of Peano Arithmetic. Oxford University Press, Oxford (1991)
11. Kotlarski, H.: Bounded induction and satisfaction classes. Zeitschrift für matematische Logik und Grundlagen der Mathematik **32**, 531–544 (1986)
12. Kotlarski, H., Krajewski, S., Lachlan, A.: Construction of satisfaction classes for nonstandard models. Can. Math. Bull. **24**, 283–293 (1981)

13. Krajewski, S.: Nonstandard satisfaction classes. In: Marek, W., et al. Set Theory and Hierarchy Theory: A Memorial Tribute to Andrzej Mostowski. Lecture Notes in Mathematics, vol. 537, pp. 121–144 . Springer, Berlin (1976)
14. Leigh, G.: Conservativity for theories of compositional truth via cut elimination. J. Symb. Log. **80**, 845–865 (2015)
15. Łełyk, M.: Axiomatic theories of truth, bounded induction and reflection principles. Ph.D. disseration, University of Warsaw (2017)
16. Łełyk, M., Wcisło, B.: Notes on bounded induction for the compositional truth predicate. Rev. Symb. Logic **10**, 455–480 (2017)
17. Löb, M.H.: Solution of a problem of Leon Henkin. J. Symb. Logic **20**, 115–118 (1955)
18. Pakhomov, F., Walsh, J.: Reflection ranks and ordinal analysis (2018). arXiv:1805.02095
19. Pudlák, P.: The lengths of proofs. In: Buss, S. (ed.) Handbook of Proof Theory, pp. 547–637. Elsevier, Amsterdam (1998)
20. Visser, A.: Semantics and the liar paradox. In: Gabbay, D., Günthner, F. (eds.) Handbook of Philosophical Logic, vol. 4, pp. 149–240. Reidel, Dordrecht (1989)
21. Visser, A.: From Tarski to Gödel. Or, how to derive the second incompleteness theorem from the undefinability of truth without self-reference (2018). arXiv:1803.03937

Affiliations

Ali Enayat[1] · **Fedor Pakhomov**[2]

Fedor Pakhomov
pakhfn@mi.ras.ru

1 Department of Philosophy, Linguistics, and the Theory of Science, University of Gothenburg, Gothenburg, Sweden

2 Steklov Mathematical Institute of Russian Academy of Sciences, Moscow, Russia