THE REVIEW OF SYMBOLIC LOGIC
Volume 7, Number 2, June 2014

TRUTH AND SPEED-UP

MARTIN FISCHER
Mathematical Center for Mathematical Philosophy,
Ludwig-Maximilians-University Munich

Abstract. In this paper, we investigate the phenomenon of **speed-up** in the context of theories of truth. We focus on axiomatic theories of truth extending Peano arithmetic. We are particularly interested on whether conservative extensions of PA have speed-up and on how this relates to a deflationist account. We show that disquotational theories have no significant speed-up, in contrast to some compositional theories, and we briefly assess the philosophical implications of these results.

§1. Introduction. Axiomatic theories of truth have become increasingly popular, and there is a wide variety of different theories.[1] There are also different ways of comparing and categorizing the theories . One way of comparison is via their proof-theoretic strength. In the realm of truth theories based on an arithmetical theory of syntax, this is understood as the number of arithmetical theorems that are provable. One of the negative aspects of comparing the proof-theoretic strength is that it is coarse-grained. For example, all conservative extensions of the base theory will be taken on a par. A more fine-grained way of comparison is via interpretability relations.[2]

In this paper, we will take a different route for comparing an axiomatic theory of truth with its base theory. Besides the questions of conservativity and interpretability, one can investigate whether the introduction of a truth predicate allows for a simplification of proofs. This is a difficult question in general, but there is at least one aspect of simplification that allows for a formal investigation, namely the length of proofs. It is possible to investigate whether theories of truth allow for shorter proofs of arithmetical theorems.

The investigation of length of proofs started with Gödel (1936). In his short paper, Gödel states that there is arbitrary speed-up in cases of arithmetic if you go from a logic of order n to order $n + 1$.[3] A very famous case of speed-up results concerns proof systems: although a first-order sequent calculus allows for cut elimination there is a price to pay: without a cut rule, there are theorems for which the proofs are significantly longer than in a sequent calculus with cut.[4] Most interesting for our purpose are speed-up theorems established by Solovay, showing that even in the case of conservative extensions a significant speed-up is

Received: September 19, 2013.

[1] For an overview, consult Halbach (2011).

[2] A relevant interpretability relation is relative interpretability going back to Tarski *et al.* (1953). Relative interpretability is based on a translation that preserves the logical structure of formulas. Moreover, theorems of the interpreted theory are translated into theorems of the interpreting theory. A promising variant in truth contexts is the notion of truth definability suggested by Fujimoto (2010).

[3] Gödel's paper does not contain a proof of the claim. A proof can be found in Buss (1994).

[4] Compare Boolos (1984) for a philosophical argument.

© Association for Symbolic Logic, 2014

doi:10.1017/S1755020314000070

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

possible. For example, the class theory NBG has speed-up over ZF. Also, the second-order arithmetic ACA_0 has speed-up over PA.

The question of speed-up for axiomatic theories of truth is not only of interest because it allows a fine-grained comparison of different theories, especially for conservative extensions of Peano arithmetic. In addition to its attractiveness as a purely technical result, it is philosophically well motivated and of particular relevance. In the philosophical debate on truth, one of the questions concerns the function and utility of the truth predicate. Deflationism is sometimes associated with the turn from asking what the nature of truth is to the question what the function of the truth predicate is.[5] The answer deflationism gives is that the only function the truth predicate has is an expressive function.[6] Although this might sound very attractive at first sight it is not at all clear what 'expressive function' exactly means, and in my opinion this is one of the blind spots in the debate on deflationism. Often in the debate the expressive function is illustrated by claiming that the truth predicate allows for generalizations.[7] Although generalizations are part of the story, there is no clear criterion available at the moment for expressiveness. We think that speed-up can play a role in explicating the utility and expressive function of the truth predicate. Simplification of proofs is connected to the expressive function of the truth predicate. We will see in the proof of our speed-up result that the main reason for a shortening of proofs is the provability of a generalization that we cannot prove in the base theory. In the more interesting case, that is, in the case of a conservative extension, the relevant generalization cannot even be formulated in the language of the base theory. We need the extra expressive resources of the truth predicate. This connection suggests a clear sufficient criterion for the increase of expressive power, namely significant speed-up of a theory of truth over its base theory. However we do not want to claim that speed-up exhausts the expressive function of the truth predicate.

The claim that the truth predicate has an expressive function is not the main point that separates deflationism from other conceptions of truth. Although deflationism emphasizes the expressive function, most other conceptions are well compatible with it. The aspect that sets deflationist theories apart is the conception that one should not inflate the notion of truth. This conception is already implicit in the claim that the expressive function is the only function of the truth predicate. Sometimes this is expressed by saying that deflationist truth is not substantial.[8] The notion of nonsubstantiality itself does not clarify much, but a clear criterion has been suggested by Horsten (1995), Shapiro (1998), and Ketland (1999): Only conservative extensions of the base theory are acceptable for a deflationist. We do not want to contribute to the ongoing debate whether deflationist theories are committed to conservative extensions or not, but just mention that for those who accept the criterion it can be an additional philosophical motivation to investigate theories of truth that are conservative extensions of Peano arithmetic.

---

[5] See, for example, Horsten (2011).

[6] Horwich, for example, writes:

> The minimalist picture of truth has three principal components: first, an account of the utility of truth (namely, to enable the explicit formulation of schematic generalizations); ... (Horwich, 1998, p. 145)

[7] Quine (1970) was the first to emphasize this role of the truth predicate.

[8] Compare Horsten (2011, p. 79).

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

In order to evaluate relevant theories of truth, we will restrict ourselves to axiomatic theories of truth formulated in the language $\mathcal{L}_{\mathsf{T}}$, which is an expansion of the arithmetical language $\mathcal{L}_{\mathsf{A}}$ by a one-place truth predicate $\mathsf{T}(x)$. Moreover, we consider only extensions of Peano arithmetic, PA. The connection of simplification of proofs and increase of expressive power gives us a philosophical motivation for the main question we want to address in this paper: *Does a theory of truth T extending* PA *have significant speed-up over* PA *with respect to* PA *theorems?*

The paper is structured as follows. In the second section, we introduce the basic notions in order to give an adequate definition of speed-up. In the third section, we show that for disquotational theories of truth the answer to our main question is negative. In the fourth section, we will introduce the basic concepts that will be necessary to characterize the features of a theory of truth for which the answer to our main question will be positive. Such a theory will then be presented in the fifth section. A compositional theory of positive truth, PT^{-}, is chosen as an example for significant speed-up. This is followed by a discussion of the case of classical compositional truth in the sixth section. In the last section, we will discuss the consequences of these results for a deflationist account of truth.

§2. Basics. The question of speed-up depends on several choices. One choice is the underlying calculus that is used. Another choice is which measure is applied. A third choice depends on what counts as a significant speed-up.

We are going to use a schematic Hilbert style axiomatic system based on the rules of modus ponens and generalization.[9] There are two measures of the length of a proof: the symbol-length and the step-length. For a proof in a Hilbert style calculus d, the symbol-length of d is the total number of occurrences of symbols in d and the step-length of d is the total number of occurrences of formulas in d. In the following we will restrict ourselves to the symbol-length for several reasons. First of all, counting all the symbols is more truthful to the task. Every last small difference in the formulation gets counted, whereas in the case of step-length very long formulas can be used in a single step. Moreover, some of the speed-up results are established only for the case of symbol-length and it is not known whether those results can be transferred to the case of step-length. To be more precise we will use the notion of size of a proof that corresponds to the symbol-length.

We consider a standard encoding of formulas and proofs as sequences in a finite alphabet.

DEFINITION 2.1. *The* ***size of a formula*** φ *is the length of the sequence coding* φ *and will be denoted by* $|\varphi|$. *The* ***size of a proof*** d *is the length of the code for* d *and will be denoted by* $|d|$.

We also write $d : A \vdash \varphi$ for d is a proof of φ from A in our Hilbert-style calculus, where A is a set of axioms of a theory. And $A \vdash \varphi$ for φ is derivable in A, and $A \vdash^{n} \varphi$ for: there is a proof of φ in A of size $\leq n$. Moreover we use

$$||\varphi||_A = \begin{cases} \text{minimal } n \text{ such that } A \vdash^{n} \varphi, \text{ if } A \vdash \varphi \\ \infty \text{ otherwise} \end{cases}$$

It is also convenient to use the big-O or Landau notation from complexity theory.

[9] For a presentation of such a system, consult Pudlák (1998, p. 552f).

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

DEFINITION 2.2. *Let f, g be functions: $\mathbb{N} \to \mathbb{N}$, then*

$$f(x) = O(g(x)) \text{ iff there is a } x_0 \text{ and } k, \text{ such that for all } x > x_0,\ f(x) \leq k \cdot g(x).$$

What does it mean for a speed-up to be significant? Two important classes of functions are the polynomial and the (Kalmar) elementary functions. In the case of a difference that is only polynomial, we are reluctant to say that it is significant. In complexity theory, it is very common to neglect differences of polynomial size.[10]

The class of functions with Kalmar elementary growth rate contains the polynomial functions. A natural suggestion for a sufficient criterion for significant increase is the first function that majorizes all the Kalmar elementary functions. The superexponential function is the first function in the Wainer-hierarchy that majorizes all Kalmar elementary functions. The general formulation is 2^n_m defined by recursion as follows: $2^n_0 = n$ and $2^n_{m+1} = 2^{2^n_m}$. We often use the special case 2^0_m and just use the abbreviation 2_m in this case. We can then characterize the Kalmar elementary functions as follows: a function f has Kalmar elementary growth rate if there is a natural number m, such that $f(x)$ is eventually majorized by 2^x_m. This allows for a precise definition of nonelementary speed-up:[11]

DEFINITION 2.3 (Speed-up). *Let T, T' be two recursive sets of axioms and $\overline{T}$ and $\overline{T'}$ the corresponding sets of theorems, such that the signature of T is a subset of the signature of T' and $\overline{T} \subseteq \overline{T'}$. Let $\Sigma \subseteq \overline{T}$.*

(i). *T' has at most polynomial speed-up over T if there exists a polynomial P with natural coefficients, such that for all φ with $T \vdash \varphi$, then*

$$||\varphi||_T < P(||\varphi||_{T'}).$$

(ii). *T' has nonelementary speed-up over T with respect to Σ if there is a sequence of formulas $\{\varphi_i\}_{i\in\omega} \subseteq \Sigma$ such that there is no function f with Kalmar elementary growth rate satisfying $||\varphi_i||_T < f(||\varphi_i||_{T'})$ for all $i \in \omega$.*

All the examples mentioned in the introduction are examples of nonelementary speed-up.

§3. Disquotational truth. In this section, we investigate the length of proofs for disquotational theories of truth, that is, theories that take the T-sentences as truth theoretic axioms. The arithmetical language $\mathcal{L}_\mathsf{A}$ contains $(\mathsf{0}, \mathsf{S}, +, \times, =)$ and $\mathcal{L}_\mathsf{T}$ is $\mathcal{L}_\mathsf{A}$ expanded by a one-place predicate T. In order to state the theory we follow some notational conventions:

We assume some appropriate fixed standard Gödel-numbering. If e is an $\mathcal{L}_A$-expression, then we will use $\#e$ to refer to the Gödelnumber of e, whereas $\ulcorner e \urcorner$ is the canonical term, that is, the numeral, representing the Gödelnumber of e in $\mathcal{L}_\mathsf{A}$. The usual definition of numerals $\underline{n}$ as $\mathsf{S}...\mathsf{S}(0)$ is not suitable for our purposes. For a number n given in binary presentation, that is $n = \sum_{i=0}^{k} 2^i a_i$, $a_i \in \{0, 1\}$, we take $\underline{n} = \underline{a_0} + \underline{2} \cdot (\underline{a_1} + 2 \cdot (\underline{a_2} + ... (\underline{a_{k-1}} + 2 \cdot \underline{a_k})...))$, where $\underline{1} = \mathsf{S}(0)$ and $\underline{2} = \underline{1} + \underline{1}$. This gives us that $|\underline{n}|$ is $O(\mathsf{log}\ n)$.[12] We also assume that our language contains function symbols for some primitive recursive functions. For example, the NUM function that sends a number n to the Gödelnumber of its numeral that is $\#\underline{n}$ is represented by num. In some cases, we also write a dot $\dot{}$ above its argument to

[10] Compare Caldon & Ignjatović (2005, p. 6).
[11] Compare, for example, Caldon & Ignjatović (2005).
[12] See also Pudlák (1998, p. 580).

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

represent the NUM function, so $\dot{x}$ is the same as $\mathsf{num}(x)$. We also have the p.r. substitution function $\mathrm{SUB}(n, m, k)$ which is $\#\varphi(t/v)$, that is, the Gödelnumber of a formula in which all occurrences of the variable v have been replaced by the term t, in the case that $n = \#\varphi$, $m = \#v$, $k = \#t$. We will use sub as a representation of this substitution function. In the case that $\varphi(v_1, \ldots, v_n)$ is a formula with at most the displayed variables occurring free in it, then we will use $\ulcorner\varphi(\dot{x}_1, \ldots, \dot{x}_n)\urcorner$ as shorthand for $\mathsf{sub}(\ldots\mathsf{sub}(\ulcorner\varphi\urcorner, \ulcorner v_1\urcorner, \dot{x}_1)\ldots, \ulcorner v_n\urcorner, \dot{x}_n)$.

We first consider the simplest case $\mathsf{TB}\restriction$, which is PA extended by the T-sentences for arithmetic, that is, for all sentences $\varphi \in \mathcal{L}_\mathsf{A}$ the instance of the following scheme:

$$(\mathsf{TB}_\varphi) \qquad \mathsf{T}(\ulcorner\varphi\urcorner) \leftrightarrow \varphi.$$

We give an outline of a proof that $\mathsf{TB}\restriction$ has at most polynomial speed-up over PA. Pudlák (1998) has proved all the requirements necessary to obtain a stronger result namely for UTB, a stronger version of disquotationalism. UTB is PA extended by instances of induction for all formulas of $\mathcal{L}_T$ and all uniform T-sentences for arithmetic, that is, for all formulas $\varphi(x_1, \ldots, x_k) \in \mathcal{L}_A$ the instance of the following scheme:

$$(\mathsf{UTB}_\varphi) \qquad \forall x_1 \ldots \forall x_k\, (\mathsf{T}(\ulcorner\varphi(\dot{x}_1, \ldots, \dot{x}_k)\urcorner) \leftrightarrow \varphi(x_1, \ldots, x_k)).$$

Both theories, TB and UTB, are conservative extensions of PA.

One way to prove the conservativity of $\mathsf{TB}\restriction$ over PA is to give a constructive method of how to transform a proof of an arithmetical formula φ in $\mathsf{TB}\restriction$ into a proof of φ in PA. The basic idea is to replace the truth predicate in the proof by an arithmetically definable partial truth predicate. Let $\varphi_1, \ldots, \varphi_n$ be an enumeration of finitely many different sentences of $\mathcal{L}_\mathsf{A}$, that is $\varphi_i \neq \varphi_j$ for all $i \neq j$ with $1 \leq i, j \leq n$. Then we can define a partial truth predicate $\mathsf{T_n}$ for these formulas in $\mathcal{L}_A$, such that for all $1 \leq i \leq n$, $\mathsf{PA} \vdash \mathsf{T_n}(\ulcorner\varphi_i\urcorner) \leftrightarrow \varphi_i$. Take

$$\mathsf{T_n}(x) := (x = \ulcorner\varphi_1\urcorner \wedge \varphi_1) \vee \ldots \vee (x = \ulcorner\varphi_n\urcorner \wedge \varphi_n).$$

This transformation gives a local interpretation of $\mathsf{TB}\restriction$ in PA as only finitely many instances of TB_φ are used in a proof. To get the desired result on the length of proofs, we have to give a version of this local interpretation that does not increase the size of proofs too much.

For an adequate estimate on the length of proofs we will follow Pudlák's presentation in Pudlák (1998). We will use a theorem by Pudlák that shows that it is possible to define partial truth predicates in a polynomial way. For a set Δ of arithmetical formulas with restricted complexity we can define a truth predicate T_Δ. There are different ways to restrict the complexity of the formulas. We are going to assume an upper bound n on the depth of the formulas in Δ.

THEOREM 3.1 (Pudlák).[13] *There is a sequence of formulas* $\mathsf{Sat}_n(x, y)$ *of polynomial size* (*in* n) *definable in* $\mathcal{L}_\mathsf{A}$, *such that for every* $\varphi(y_1, \ldots, y_k)$ *with depth* $\leq n$, *the theory* PA *proves*

$$\mathsf{Sat}_n(\ulcorner\varphi\urcorner, x) \leftrightarrow \varphi((x)_1, \ldots, (x)_k).$$

using a proof of polynomial size (*in* n).

As a consequence we get the following for a theory of universal Tarski biconditionals formulated with a satisfaction predicate. Let $\mathsf{UTB_{sat}}$ be the theory formulated in the language

[13] For a proof see Pudlák (1998, p. 560f, Theorem 3.3.1, p. 562).

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

of $\mathcal{L}_{\mathsf{sat}}$, which is $\mathcal{L}_{\mathsf{A}}$ expanded by a two place predicate Sat. The theory $\mathsf{UTB}_{\mathsf{sat}}$ contains PA, all instances of the induction scheme for the language $\mathcal{L}_{\mathsf{sat}}$ and for all formulas $\varphi(y_1, \ldots, y_k)$

$$(\mathsf{UTB}^{\mathsf{sat}}_{\varphi}) \qquad \forall x\, (\mathsf{Sat}(\ulcorner \varphi \urcorner, x) \leftrightarrow \varphi((x)_1, \ldots, (x)_k)).$$

THEOREM 3.2. $\mathsf{UTB}_{\mathsf{sat}}$ *has at most polynomial speed-up over* PA.

Proof. By Theorem 3.1 there are polynomials P_1, P_2 such that for all n, $|\mathsf{Sat}_n| < P_1(n)$ and for all φ with depth less or equal to n, there is a proof $d : \mathsf{PA} \vdash \mathsf{Sat}_n(\ulcorner \varphi \urcorner, y) \leftrightarrow \varphi((y)_1, \ldots, (y)_k)$ with $|d| < P_2(n)$.

Let $e : \mathsf{UTB}_{\mathsf{sat}} \vdash^n \varphi$ with $\varphi \in \mathcal{L}_A$. We will transform e into a proof e' in PA. In the first step we replace every occurrence of Sat in e by Sat_n. This will be sufficient, as the depth of all formulas ψ used in e is smaller than n. The result e^* is a sequence of formulas of $\mathcal{L}_{\mathsf{A}}$ and since the formula Sat_n is polynomial e^* we can give an upper bound of $|e^*| < n \cdot P_1(n)$. But e^* is not a proof. Induction axioms used in e containing the satisfaction predicate are turned into arithmetical induction axioms. But instances of $\mathsf{UTB}^{\mathsf{sat}}_{\varphi}$ are not translated into axioms, but into PA-provable formulas. To turn e^* into a proof e' we have to replace every translation of an instance ψ of $\mathsf{UTB}^{\mathsf{sat}}_{\varphi}$ by a proof e_ψ of this translation. But we know that there are such proofs e_ψ, such that $|e_\psi| < P_2(n)$. And since n is greater than the number of instances of $\mathsf{UTB}^{\mathsf{sat}}_{\varphi}$ used in e we can give an upper bound: $|e'| \leq n \cdot P_2(|e^*|) \leq n \cdot P_2(n \cdot P_1(n))$, which is polynomial in n. □

The theory $\mathsf{UTB}_{\mathsf{sat}}$ is a more general version of UTB. It is easy to define a truth predicate from a satisfaction predicate, that is, we can give a definitional extension of $\mathsf{UTB}_{\mathsf{sat}}$ that contains UTB. The definitional extension clearly preserves the arithmetical vocabulary from which we can infer that also UTB and TB have at most polynomial speed-up over PA.

Regarding disquotational theories of truth we have to draw the following conclusion. Typed disquotational theories of truth, that is, theories that contain only T-biconditionals for arithmetical sentences, have no significant speed-up over PA, and in this respect the introduction of a disquotational truth predicate does not increase our expressive ressources. This does not answer the question for type-free disquotational theories of truth, that is, disquotational theories that allow also T-biconditionals for sentences containing the truth predicate, such as PUTB. We will see in the next section that for type-free disquotational theories that prove the consistency of PA we get nonelementary speed-up by a very general result. However, a more interesting question is whether there are 'natural' conservative type-free disquotational theories that simplify proofs in a nontrivial way.[14] We will not pursue this question here but leave it for future research.

§4. Nonelementary speed-up. In contrast to the disquotational theories of truth there are typed compositional theories of truth that have nonelementary speed-up over PA. Before we give some examples, we introduce the relevant concepts used in the proofs. The results of this section are known and are repeated here for expository purposes. The method of proof we are going to use relies heavily on a method developed by Solovay. Solovay's original proof was not published and therefore we will mainly follow the Pudlák (1998) presentation. Moreover, we try to indicate the persons who proved the theorems first. Even

[14] The restriction 'natural' is supposed to exclude artificial type-free disquotational theories, for example those construed with McGee's trick. Compare McGee (1992).

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

if we are not able do this for all of them, at least we do not claim any originality of results in this section.

For a theory T to have nonelementary speed-up over PA, it is sufficient that there is a sequence of formulas $\varphi_0, \varphi_1, \ldots$, provable in PA, for which the following two claims hold:

- there is no function f with Kalmar elementary growth rate, such that for all n, $||\varphi_n||_{\mathsf{PA}} = f(n)$;
- there is a polynomial function g, such that for all n, $||\varphi_n||_T = g(n)$.

The first claim is a lower bound on proofs of those φ_i in PA and the second is an upper bound on proofs of the φ_i in T. The formulas used to establish the speed-up result are partial consistency statements. A theorem by Friedman and Pudlák shows that there are no short proofs for those partial consistency statements in PA. This allows for the possibility of partial consistency statements for which there is no function with Kalmar elementary growth rate, which is a bound on the length of proofs of those statements in PA. Those partial consistency statements will be used to prove the first claim.

To prove the second claim, one has to show that there are short proofs of those partial consistency statements in a specific theory T, that is, the length of proofs can be bound by a polynomial. This can either be established by a consistency proof for PA in T or more interestingly, for conservative theories, it is sometimes possible to define a cut and prove the consistency of PA on this cut. In this section we are going to repeat basic results concerning partial consistency statements and definable cuts.

Let $\mathsf{Pr}_{\mathsf{pa}}(\underline{n}, \ulcorner\varphi\urcorner)$ be a polynomial representation of $||\varphi||_{\mathsf{PA}} \leq n$, such that whenever it is true that $||\varphi||_{\mathsf{PA}} \leq n$, then $\mathsf{Pr}_{\mathsf{pa}}(\underline{n}, \ulcorner\varphi\urcorner)$ has a proof of polynomial length in n in Q.[15] Let $\mathsf{Con}_{\mathsf{pa}}(\underline{n})$ be the statement of consistency of PA up to length $\leq n$, that is

$$\mathsf{Con}_{\mathsf{pa}}(x) := \neg\mathsf{Pr}_{\mathsf{pa}}(x, \ulcorner 0 = 1\urcorner)$$

These partial consistency statements are all Π_1^b.

THEOREM 4.1 (Friedman/Pudlák).[16] *Let T be a sufficiently strong fragment of arithmetic, axiomatized by an $\mathcal{NP}$ set of axioms. Then there exists an $\varepsilon > 0$, such that for all n,*

$$||\mathsf{Con}_T(\underline{n})||_T > n^{\varepsilon}.$$

In order to connect this with the superexponentiation function we use the following theorem.[17]

THEOREM 4.2. *Let T be a sufficiently strong theory. Let f be a provably total function in T whose graph has a polynomial numeration in T. Then there exists an $\varepsilon > 0$ such that*

$$||\mathsf{Con}_T(f(\underline{n}))||_T \geq f(n)^{\varepsilon},$$

while

$$||\mathsf{Con}_T(f(\underline{n}))||_{T+\mathsf{Con}_T} = O(\log n).$$

In the case that T is strong enough to formalize the following lemma then the restriction to functions whose graph has a polynomial numeration is not essential:[18]

---

15 See Pudlák (1998, p. 580).

16 Pudlák (1998, p. 581, Theorem 6.2.3) or Buss (1994, p. 741, Theorem 4).

17 See Pudlák (1998, p. 588).

18 See Pudlák (1998, p. 587).

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

LEMMA 4.3. *For any recursive function f, there exists a recursive function g such that*

- $\forall n \in \mathbb{N}\ f(n) \leq g(n)$,
- *the graph of g is a polynomial time computable function.*

Important for our purposes is the following corollary:

COROLLARY 4.4. $||\mathsf{Con}_{\mathsf{pa}}(2_{\underline{n}})||_{\mathsf{PA}} \geq (2_n)^{\epsilon}$, *for some constant* $\epsilon > 0$.

The full consistency statement can be understood as a generalization of all the partial consistency statements and we can easily derive all the partial consistency statements from it in a direct and short manner. This gives us the following corollary:[19]

COROLLARY 4.5. $\mathsf{PA} + \mathsf{Con}_{\mathsf{pa}}$ *has nonelementary speed-up over* PA.

This corollary shows that for all axiomatic theories of truth that contain PA and prove the consistency of PA we have nonelementary speed-up. Since all theories of truth that prove the consistency of PA are on a par with regard to speed-up, we have a reason to focus on conservative theories of truth as in this case the question is not already answered by this corollary. Also from another perspective we have a reason to focus on conservative extensions. The consistency statement for PA is a generalization that we can state in $\mathcal{L}_{\mathsf{A}}$. In this case we cannot fully separate the deductive and expressive power of truth. Here the reason for speed-up relies on the ability to prove a generalization, that is expressible without the truth predicate. In contrast, conservative theories of truth can only prove generalizations, which are either also provable in PA or which we cannot formulate without the truth predicate. It is the second form of generalizations that we are interested in, because these are the generalizations that are signs of increase of expressive power. One of those generalizations will allow us to derive all the partial consistency statements in a short manner.

In order to formulate such a generalization, the notion of a cut is helpful. In the model theory of PA, cuts are initial segments of a nonstandard model, that is, sets that contain 0 are closed under the successor function S and if n is in the set then also all m that are considered by the model to be smaller than n are in the set. The natural numbers are a cut in this sense in any nonstandard model of PA. A definable cut is a formula that defines in any model an initial segment. In the following we will only talk about definable cuts.[20] They play an important role in the study of speed-up and will be used frequently later on. In the following we assume that $T \supseteq \mathsf{Q}$ is a sequential theory, that is, a theory that is capable of coding finite sequences, in a language expansion of $\mathcal{L}_{\mathsf{A}}$.

A formula $\mathsf{I}(x)$ is called inductive in T iff the following statements are provable:

(i). $T \vdash \mathsf{I}(0)$.

(ii). $T \vdash \forall x\ (\mathsf{I}(x) \rightarrow \mathsf{I}(\mathsf{S}x))$.

A formula $\mathsf{C}(x)$ is a cut in T iff it is inductive and

(iii). $T \vdash \forall x \forall y\ (\mathsf{C}(x) \wedge y < x \rightarrow \mathsf{C}(y))$.

A cut $\mathsf{C}(x)$ is called proper if $T \nvdash \forall x\, \mathsf{C}(x)$.

---

[19] Stronger results are known. Basically if T is a sufficiently strong theory of arithmetic that proves $\mathsf{Con}_{\mathsf{pa}}$, then T has speed-up by any provably total recursive function of T. Compare Buss (1994).

[20] The material on cuts presented in this paper is well known. The method of shortening cuts was developed by Solovay (unpublished). For more on cuts, see for example Hájek & Pudlák (1993).

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

We clearly have that for all cuts $\mathsf{C}(x)$ in T and all $n \in \mathbb{N}$ we have $T \vdash \mathsf{C}(\underline{n})$. Cuts already have some nice properties; however, there are ways to get cuts with more interesting properties. Solovay developed a powerful method of shortening cuts. The first thing to notice is that every inductive formula gives rise to a cut. Inductive formulas may be satisfied by two numbers n, m and not by k with $n < k < m$. However, there is an initial segment in which all elements satisfy the original inductive formula and we can define this initial segment by shortening.

LEMMA 4.6. *Let* $\mathsf{I}(x)$ *be an inductive formula in* T. *Then there is a cut* $\mathsf{C}_{\mathsf{I}}(x)$ *in* T, *such that* $T \vdash \forall x\,(\mathsf{C}(x) \to \mathsf{I}(x))$.

Proof. The idea of the proof is to restrict the inductive formula $\mathsf{I}(x)$. So let

$$\mathsf{C}(x) :\leftrightarrow \forall y\,(y \leq x \to \mathsf{I}(y)).$$

Then it is straightforward to check that $\mathsf{C}(x)$ is a cut and $\mathsf{C}(x) \to \mathsf{I}(x)$.[21] □

Moreover, every cut can be shortened to a cut that is closed under addition and multiplication.

LEMMA 4.7. *Let* $\mathsf{C}(x)$ *be a cut in* T. *Then there is a formula* $\mathsf{C}^+(x)$ *such that*

(i). $\mathsf{C}^+(x)$ *is a cut in* T;
(ii). $T \vdash \forall x\,(\mathsf{C}^+(x) \to \mathsf{C}(x))$;
(iii). $T \vdash \forall x \forall y\,(\mathsf{C}^+(x) \wedge \mathsf{C}^+(y) \to \mathsf{C}^+(x + y))$.

Proof. For $\mathsf{C}(x)$ a cut in T we can define $\mathsf{C}^+(x) := \forall z\,(\mathsf{C}(z) \to \mathsf{C}(z + x))$.
$\mathsf{C}^+(0)$ we get directly. $\mathsf{C}^+(x) \to \mathsf{C}^+(\mathsf{S}x)$ by the fact that $T \vdash \mathsf{S}(z + x) = z + \mathsf{S}x$. So C^+ is inductive in T. Since C is a cut C^+ is also a cut. It is straightforward to check that (ii) is satisfied. (iii) Holds because of the associativity of addition. □

With a similar construction we get a cut that is closed under addition and multiplication.

LEMMA 4.8. *Let* $\mathsf{C}^+(x)$ *be a cut in* T *closed under addition. Then there is a formula* $\mathsf{C}^\times(x)$ *such that*

(i). $\mathsf{C}^\times(x)$ *is a cut in* T;
(ii). $T \vdash \forall x\,(\mathsf{C}^\times(x) \to \mathsf{C}^+(x))$;
(iii). $T \vdash \forall x \forall y\,(\mathsf{C}^\times(x) \wedge \mathsf{C}^\times(y) \to \mathsf{C}^\times(x \times y))$.
(iv). $T \vdash \forall x \forall y\,(\mathsf{C}^\times(x) \wedge \mathsf{C}^\times(y) \to \mathsf{C}^\times(x + y))$.

Proof. Take $\mathsf{C}^\times(x) :\leftrightarrow \forall z\,(\mathsf{C}^+(z) \to \mathsf{C}^+(z \times x))$. That $\mathsf{C}^\times(x)$ is a cut is proved similar to the previous proof. (ii) is also checked directly. For (iii) we rely on the fact that we can prove the associativity of multiplication in T. For (iv) $\mathsf{C}^\times(x) \to \forall z\,(\mathsf{C}^+(z) \to \mathsf{C}^+(z \times x))$ and $\mathsf{C}^\times(y) \to \forall z\,(\mathsf{C}^+(z) \to \mathsf{C}^+(z \times y))$. Since C^+ is closed under addition $\mathsf{C}^\times(x) \wedge \mathsf{C}^\times(y) \to \forall z\,(\mathsf{C}^+(z) \to \mathsf{C}^+(z \times x + z \times y))$ and therefore $\mathsf{C}^\times(x) \wedge \mathsf{C}^\times(y) \to \forall z\,(\mathsf{C}^+(z) \to \mathsf{C}^+(z \times (x + y))$. □

[21] This works for theories for which we can prove the transitivity of $<$, but not in the case of Q. In this case we have to restrict ourselves to a transitively closed inductive formula. In this case let $\psi_<(x)$ be the formula $\forall y \forall z\,(z < y \wedge y < x \to z < x)$. It is easy to check that $\psi_<(x)$ is inductive. Further we know that the conjunction of two inductive formulas is again an inductive formula. So let $\mathsf{I}_<(x) := \mathsf{I}(x) \wedge \psi_<(x)$ and $\mathsf{C}(x) := \forall y\,(y \leq x \to \mathsf{I}_<(y))$.

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

What we do not get in general is a cut that is closed under exponentiation. This method of shortening cuts doesn't work for the reason that exponentiation is not associative.[22] However, for every cut $\mathsf{C}(x)$ in S we can define a new cut $\mathsf{C}^*(x) :\leftrightarrow \mathsf{C}(2^x)$ such that $S \vdash \forall x\,(\mathsf{C}^*(x) \rightarrow \mathsf{C}(x) \wedge \mathsf{C}(2^x))$. This construction can be iterated. Moreover, this can be achieved in a schematic way such that the proofs are polynomial in n.[23] In the following let $\mathsf{Cut}(\varphi)$ be short for the claim that φ is a cut formula, that is $\varphi(0) \wedge \forall x\,(\varphi(x) \rightarrow \varphi(\mathsf{S}x)) \wedge \forall x \forall y\,(\mathsf{C}(x) \wedge y < x \rightarrow \mathsf{C}(y))$.

THEOREM 4.9. *Let T be a theory containing a sufficiently strong fragment of arithmetic. Let $\varphi_0(x)$ be a cut in T. Then there is a sequence of formulas $\varphi_1, \varphi_2, \ldots$ such that for each n the following formula*

$$\mathsf{Cut}(\varphi_{n+1}(x)) \wedge \forall x\,(\varphi_{n+1}(x) \rightarrow (\varphi_n(x) \wedge \varphi_n(2^x)))$$

has a proof in T of size polynomial in n.

Combining the proofs for $\varphi_n(0) \rightarrow \varphi_{n-1}(2^0)$, $\varphi_{n-1}(2^0) \rightarrow \varphi_{n-2}(2^{2^0}), \ldots, \varphi_1(2_{\underline{n-1}}) \rightarrow \varphi_0(2_{\underline{n}})$ we get a proof of $\varphi_0(2_{\underline{n}})$ that is polynomial in n.

COROLLARY 4.10. *Let T be a theory containing a sufficiently strong fragment of arithmetic. Let $\mathsf{C}(x)$ be a cut in T. Then $||\mathsf{C}(2_{\underline{n}})||_T = n^{O(1)}$.*

Cuts are themselves very interesting and have important consequences. For example Pudlák uses the method of cuts to give a strengthening of Gödel's second incompleteness theorem that says that no consistent theory containing Q proves its consistency on a cut.[24] Cuts are also used to show the interpretability of some bounded arithmetics in Q.[25]

Combining the facts about partial consistency statements and cuts gives a way to prove speed-up results for theories that prove the consistency of PA on a cut.

THEOREM 4.11. *Let T be a theory that proves the consistency of PA on a cut, that is, there is a formula $\mathsf{C}(x)$ in $\mathcal{L}_T$ such that $\mathsf{C}(x)$ is a cut in T and $T \vdash \forall x\,(\mathsf{C}(x) \rightarrow \mathsf{Con}_{\mathsf{pa}}(x))$. Then $||\mathsf{Con}_{\mathsf{pa}}(2_{\underline{n}})||_T = n^{O(1)}$.*

Proof. Let $\mathsf{C}(x)$ be a cut in T such that $T \vdash \forall x\,(\mathsf{C}(x) \rightarrow \mathsf{Con}_{\mathsf{pa}}(x))$. By Corollary 4.10, we have $||\mathsf{C}(2_{\underline{n}})||_T = n^{O(1)}$. Combining these proofs, suffices to show that $||\mathsf{Con}_{\mathsf{pa}}(2_{\underline{n}})||_T = n^{O(1)}$. □

Combining Corollary 4.4 and Theorem 4.11 we get a sufficient criterion for speed-up:

THEOREM 4.12. *Let T be a theory that proves the consistency of PA on a cut. Then T has nonelementary speed-up over PA.*

In contrast to the statement $\mathsf{Con}_{\mathsf{pa}}$, which is a sentence of $\mathcal{L}_{\mathsf{A}}$, a statement of consistency on a cut $\forall x\,(\mathsf{C}(x) \rightarrow \mathsf{Con}_{\mathsf{pa}}(x))$ does not imply nonconservativity. The reason being that the restriction to a cut C may contain new vocabulary.

---

22 Compare Burgess (2005, p. 103f) for an elaboration of this point. Paris & Dimitracopoulos (1983) show the impossibility.

23 For a proof see Pudlák (1998, p. 563).

24 See Pudlák (1985).

25 Compare Hájek & Pudlák (1993, p. 366f).

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

This proof method has been used by Solovay to establish speed-up results for ACA_0 over PA as well as GB over ZF. So for example ACA_0 proves the consistency of PA on a cut, which allows one to prove that ACA_0 has superexponential speed-up over PA.

§5. Positive compositional truth. We are now ready for a first application to compositional theories of truth. In the next subsection we are going to show that the theory of truth PT^- has nonelementary speed-up over PA. In order to state the theory we follow some conventions:

The set of Gödelnumbers of arithmetical sentences is represented by Snt, the set of arithmetical formulas by Fml, the set of all formulas with at most one free variable by Fml^1. $\#\mathsf{FV}(x)$ stands for the number of free variables of the formula x. The set of closed terms is represented by Ct. Moreover, we will use the letters s, t with the intended reading of variables that range over closed terms, that is $\forall t\, \varphi(t)$ is an abbreviation for $\forall x\, (\mathsf{Ct}(x) \rightarrow \varphi(x))$. We represent the p-time function that applied to a closed term gives the value of that term by val. The function that takes the Gödelnumber of a formula to the Gödelnumber of its universal closure of it, is represented by ucl. For the representation of syntactical operations we use the dot notation, for example $x \dot{\wedge} y$ represents the concatenation of x and the conjunction symbol and y. In the case that x is a formula with one free variable v we just use $x(t)$ short for $\mathsf{sub}(x, \ulcorner v \urcorner, t)$. We will also use the vector notation $\vec{t}$ to refer to sequences of closed terms. So $\forall \vec{t}\, \varphi(\vec{t}\,)$ is short for $\forall x\, (\mathsf{Seq}(x) \wedge \forall i \leq \mathsf{lh}(x)\, \mathsf{Ct}((x)_i) \rightarrow \varphi(x))$. Moreover we assume some enumeration of the free variables v_i in a formula, such that for a formula x with n free variables and $\vec{t}$ a sequence of terms of length greater or equal to n, $x(\vec{t}\,)$ is short for $x((t)_1/v_1, \ldots (t)_n/v_n)$.

PT^- is then given by the universal closures of the following axioms:

(C1) $\mathsf{T}(s \doteq t) \leftrightarrow \mathsf{val}(s) = \mathsf{val}(t)$.

(C2) $\mathsf{T}(\dot{\neg}\, (s \doteq t)) \leftrightarrow \neg\, \mathsf{val}(s) = \mathsf{val}(t)$.

(C3) $\mathsf{Snt}(x \dot{\wedge} y) \rightarrow (\mathsf{T}(x \dot{\wedge} y) \leftrightarrow \mathsf{T}(x) \wedge \mathsf{T}(y))$.

(C4) $\mathsf{Snt}(\dot{\neg}\, (x \dot{\wedge} y)) \rightarrow (\mathsf{T}(\dot{\neg}\, (x \dot{\wedge} y)) \leftrightarrow \mathsf{T}(\dot{\neg}\, x) \vee \mathsf{T}(\dot{\neg}\, y))$.

(C5) $\mathsf{Snt}(x \dot{\vee} y) \rightarrow (\mathsf{T}(x \dot{\vee} y) \leftrightarrow \mathsf{T}(x) \vee \mathsf{T}(y))$.

(C6) $\mathsf{Snt}(\dot{\neg}\, (x \dot{\vee} y)) \rightarrow (\mathsf{T}(\dot{\neg}\, (x \dot{\vee} y)) \leftrightarrow \mathsf{T}(\dot{\neg}\, x) \wedge \mathsf{T}(\dot{\neg}\, y))$.

(C7) $\mathsf{Snt}(\dot{\forall} y\, x) \rightarrow (\mathsf{T}(\dot{\forall} y\, x) \leftrightarrow \forall t\, \mathsf{T}(x(t)))$.

(C8) $\mathsf{Snt}(\dot{\neg}\, \dot{\forall} y\, x) \rightarrow (\mathsf{T}(\dot{\neg}\, \dot{\forall} y\, x) \leftrightarrow \exists t\, \mathsf{T}(\dot{\neg}\, x(t)))$.

(C9) $\mathsf{Snt}(\dot{\exists} y\, x) \rightarrow (\mathsf{T}(\dot{\exists} y\, x) \leftrightarrow \exists t\, (\mathsf{T}(x(t)))$.

(C10) $\mathsf{Snt}(\dot{\neg}\, \dot{\exists} y\, x) \rightarrow (\mathsf{T}(\dot{\neg}\, \dot{\exists} y\, x) \leftrightarrow \forall t\, \mathsf{T}(\dot{\neg}\, x(t)))$.

(C11) $\mathsf{Snt}(x) \rightarrow (\mathsf{T}(\dot{\neg}\, \dot{\neg}\, x) \leftrightarrow \mathsf{T}(x))$.

(C12) $\mathsf{Fml}^1(x) \rightarrow (\mathsf{T}(x(t)) \leftrightarrow \mathsf{T}(x(\mathsf{num}(\mathsf{val}(t)))))$.

Formulas for which every substitution instance by a closed term is either true or its negation is true are called total:

$$\mathsf{tot}(x) :\Leftrightarrow \mathsf{Fml}^1(x) \wedge \forall t\, (\mathsf{T}(x(t)) \vee \mathsf{T}(\dot{\neg}\, x(t))).$$

The induction axiom of PT^- is a kind of internal induction, restricted to total formulas:

($\mathrm{I_tI}$) $\mathsf{tot}(x) \wedge \mathsf{T}(x(0)) \wedge \forall u\, (\mathsf{T}(x(\dot{u})) \rightarrow \mathsf{T}(x(\mathsf{num}(\mathsf{S}u)))) \rightarrow \forall u\, \mathsf{T}(x(\dot{u}))$.

$$\mathsf{PT}^- := \mathsf{PA} \cup \text{(C1)-(C12)} \cup (\mathrm{I_tI}).$$

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

It is well known that PT^- proves all instances of the uniform T-biconditionals, that is

$$\mathsf{PT}^- \vdash \forall t\,(\mathsf{T}(\ulcorner\varphi\urcorner(t/\ulcorner x\urcorner)) \leftrightarrow \varphi(\mathsf{val}(t))) \qquad \text{, for all formulas } \varphi(x) \in \mathcal{L}_{\mathsf{A}}.$$

The axiom (C12) guarantees that the truth value of a sentence $\varphi(t)$, that is, the result from replacing the only free variable in φ by the closed term t, only depends on the value of t.[26] It allows us to prove the principle of regularity.

LEMMA 5.1. $\mathsf{PT}^- \vdash \forall x \forall s \forall t\,(\mathsf{Fml}^1(x) \wedge \mathsf{val}(s) = \mathsf{val}(t) \rightarrow (\mathsf{T}x(s) \leftrightarrow \mathsf{T}x(t)))$.

We can define the positive grade PG of a formula by saying that literals have the positive grade 0 and if φ is of the form $\neg\neg\,\psi$, $\psi \wedge \chi$, $\psi \vee \chi$, $\neg\,(\psi \wedge \chi)$, $\neg\,(\psi \vee \chi)$, then $\mathrm{PG}(\varphi) = max\{\mathrm{PG}(\psi), \mathrm{PG}(\chi)\}+1$ and if φ is of the form $\forall x\,\psi$, $\exists x\,\psi$, $\neg\forall x\,\psi$, $\neg\exists x\,\psi$, then $\mathrm{PG}(\varphi) = \mathrm{PG}(\psi) + 1$. We use pg for the representation of PG.

In the following we are going to show that PT^- has nonelementary speed-up over PA. We will do this by showing that PT^- proves the consistency of PA on a cut. So our first step will be the definition of our cut in PT^-. The cut has some nice properties that will allow us to prove the consistency of PA on this cut in the following way: first, notice that PT^- proves that logic is true on this cut, that is, it proves that all logical axioms on this cut are true. Then we show that all arithmetical axioms are true on this cut. We can then show that the rules of inference preserve truth on this cut.

We have to make sure that our formal representations $\varphi(x)$ behave in the way expected. Since in PT^- we do not have full induction for all formulas containing the truth predicate, we will sometimes not be able to show $\forall x\,\varphi(x)$, although we can show that $\varphi(\mathsf{0})$ and $\forall x\,(\varphi(x) \rightarrow \varphi(\mathsf{S}x))$. In these cases we will restrict ourselves to an initial segment whose members all satisfy φ.

In a first step we will restrict ourselves to universal quantifier blocks for which we can still prove a generalization of (C7). The reason for this restriction is that our axiom (C7) is only for the case of one universal quantifier, although we would like to have the general form for an arbitrary number of universal quantifiers. We could introduce an axiom for any fixed finite number of free variables. However later, we want to prove that all logical axioms are true. For this we will use a representation lAx of the set of logical axioms. By overspill there will be nonstandard formulas satisfying lAx, and there will be formulas where the number of free variables is nonstandard. As we don't have full induction we cannot guarantee that a generalization of (C7) holds for all formulas. In this case we have to restrict ourselves to those formulas for which the number of quantifiers is in a specific cut.

LEMMA 5.2.

$$\forall x\,(\mathsf{Fml}(x) \wedge \#\mathsf{FV}(x) = z + 1 \rightarrow (\forall \vec{t}\,(\mathsf{lh}(\vec{t}\,) > z \rightarrow \mathsf{T}x(\vec{t}\,)) \leftrightarrow \mathsf{Tucl}(x)))$$

is inductive.

Proof. Let $\mathsf{I}_{\mathsf{u}}(z)$ be the formula in question. $\mathsf{I}_{\mathsf{u}}(\mathsf{0})$ is proved by axiom (C7). For the successor step we also use (C7). □

Let $\mathsf{C}_{\mathsf{u}}(x)$ be a cut formula resulting from shortening the formula $\mathsf{I}_{\mathsf{u}}(z)$.

In a second step we will restrict ourselves to universal quantifier blocks for which we can still prove a generalization of regularity:

[26] Compare Cantini (1989, p. 102).

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

LEMMA 5.3.

$$\forall x(\mathsf{Fml}(x) \wedge \#\mathsf{FV}(x) = z+1 \rightarrow \forall \vec{s}, \vec{t}\,(\mathsf{lh}(\vec{s}\,) > z \wedge \mathsf{lh}(\vec{t}\,) > z \wedge \forall i\,(1 \leq i \leq z+1 \rightarrow \mathsf{val}((s)_i) = \mathsf{val}((t)_i)) \rightarrow (\mathsf{T}x(\vec{s}\,) \leftrightarrow \mathsf{T}x(\vec{t}\,))))$$

is inductive.

Proof. Let $\mathsf{I}_\mathsf{v}(z)$ be the formula in question. $\mathsf{I}_\mathsf{v}(0)$ is proved by regularity. For the successor step regularity is also used. □

Let $\mathsf{C}_\mathsf{v}(x)$ be a cut formula resulting from shortening the inductive formula $\mathsf{I}_\mathsf{v}(x)$. We take $\mathsf{C}_\mathsf{r}(x) := \mathsf{C}_\mathsf{u}(x) \wedge \mathsf{C}_\mathsf{v}(x)$.

By shortening we can assume that our cut is closed under addition and multiplication and since our Gödelcoding makes no use of exponentiation we have

$$\mathsf{PT}^- \vdash \forall x\,(\mathsf{C}_\mathsf{r}(x) \wedge \mathsf{Ct}(x) \rightarrow \exists y\,(\mathsf{C}_\mathsf{r}(y) \wedge \mathsf{val}(x) = y)).$$

Next we define total and complete formulas with a restricted number of free variables: $\mathsf{tc}^{\mathsf{C}_r}(x) := \mathsf{Fml}(x) \wedge \mathsf{C}_\mathsf{r}(\#\mathsf{FV}(x)) \wedge \forall \vec{t}\,(\mathsf{lh}(\vec{t}\,) \geq \#\mathsf{FV}(x) \rightarrow (\neg \mathsf{T}x(\vec{t}\,) \leftrightarrow \mathsf{T}\dot{\neg} x(\vec{t}\,)))$ We can use this restriction on the number of free variables to define a new inductive formula $\mathsf{I}_\mathsf{t}(x)$:

$$\mathsf{I}_\mathsf{t}(x) := \forall y\,(\mathsf{Fml}(y) \wedge \mathsf{pg}(y) = x \wedge \mathsf{C}_\mathsf{r}(\#\mathsf{FV}(y)) \rightarrow \mathsf{tc}^{\mathsf{C}_\mathsf{r}}(y)).$$

LEMMA 5.4. $\mathsf{I}_\mathsf{t}(x)$ *is inductive in* PT^-.

Proof. The case $\mathsf{I}_\mathsf{t}(0)$: the provably unique readability of $\mathcal{L}_A$-formulas guarantees that we can prove in PA that

$$\forall x\,(\mathsf{Fml}(x) \wedge \mathsf{pg}(x) = 0 \rightarrow \exists y, z\,(\mathsf{Term}(y) \wedge \mathsf{Term}(z) \wedge (x = y \doteq z \vee x = y \dot{\neq} z)))$$

Moreover we can prove in PA

$$\forall y, z, \vec{t}\,(\mathsf{Term}(y) \wedge \mathsf{Term}(z) \wedge \mathsf{lh}(\vec{t}\,) \geq \#\mathsf{FV}(y \doteq x) \rightarrow (\mathsf{val}(y(\vec{t}\,)) = \mathsf{val}(z(\vec{t}\,)) \vee \mathsf{val}(y(\vec{t}\,)) \neq \mathsf{val}(z(\vec{t}\,)))).$$

With the axioms (C1) and (C2) we can then prove in PT^- that

$$\forall y, z, \vec{t}\,(\mathsf{Term}(y) \wedge \mathsf{Term}(z) \wedge \mathsf{lh}(\vec{t}\,) \geq \#\mathsf{FV}(y \doteq x) \rightarrow (\mathsf{T}(\dot{\neg}\, y \doteq x(\vec{t}\,)) \leftrightarrow \neg \mathsf{T}(y \doteq x(\vec{t}\,)) \wedge \mathsf{T}(\dot{\neg}\, y \dot{\neq} x(\vec{t}\,)) \leftrightarrow \neg \mathsf{T}(y \dot{\neq} x(\vec{t}\,)))).$$

The case $\forall x\,(\mathsf{I}_\mathsf{t}(x) \rightarrow \mathsf{I}_\mathsf{t}(\dot{\mathsf{S}}\, x))$: by the positive truth clause of PT^-. The provable unique readability guarantees that formulas of positive grade $n+1$ are built up by one of the cases taken care of by (C3) − (C11). Those axioms ensure that truth determinateness, that is being tc, is preserved by the positive build up of formulas. □

Since the positive grade of a formula is less or equal to its code we get by the method of shortening a cut $\mathsf{C}_\mathsf{t}(x)$ that contains only truth determinate formulas.

LEMMA 5.5. *There is a cut* $\mathsf{C}_\mathsf{t}(x)$ *in* PT^-*, such that*

$$\mathsf{PT}^- \vdash \forall x\,(\mathsf{Fml}(x) \wedge \mathsf{C}_\mathsf{t}(x) \rightarrow \mathsf{C}_\mathsf{r}(\#\mathsf{FV}(x)) \wedge \mathsf{tc}^{\mathsf{C}_\mathsf{r}}(x)).$$

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

We can use this cut to prove that the universal closures of logical axiom instances are true on the cut. Let lAx be a formula strongly representing the primitive recursive set of being a logical axiom in the language $\mathcal{L}_A$.[27]

LEMMA 5.6. $\mathsf{PT}^- \vdash \forall x\,(\mathsf{C_t}(x) \wedge \mathsf{lAx}(x) \rightarrow \mathsf{Tucl}(x))$.

Proof. For the propositional axioms this is done in the usual way by considering one instance of the axiom schema containing the truth predicate and using the fact that the logical axiom is total and complete. For example consider the axiom schema

$$A \rightarrow (A \vee B).$$

We need to prove that

$$\forall x\,(\mathsf{C_t}(x) \wedge \exists y, z\,(\mathsf{Fml}(y) \wedge \mathsf{Fml}(z) \wedge x = y \dot{\rightarrow} (y \dot{\vee} z) \rightarrow \mathsf{Tucl}(x))).$$

We have

$$\mathsf{T}y(\vec{t}\,) \rightarrow (\mathsf{T}y(\vec{t}\,) \vee \mathsf{T}z(\vec{t}\,))$$

as an instance of the logical axiom schema itself containing the truth predicate. By our assumption x, y, z are total and complete and we get

$$\mathsf{T}(y \dot{\rightarrow} (y \dot{\vee} z))(\vec{t}\,).$$

By the assumption that the number of free variables is in the cut we can infer by Lemma 5.2 that

$$\mathsf{Tucl}(y \dot{\rightarrow} (y \dot{\vee} z)).$$

Similar for the quantifier axioms by using regularity and the generalized version of (C7). □

Moreover, due to the fact that $\mathsf{UTB}\restriction \subseteq \mathsf{PT}^-$, we can prove in PT^- that all axioms of Q are true. What we need to prove is that all induction axioms of PA are true on this cut. Let $\mathsf{IndAx_{pa}}(x)$ be an arithmetical formula strongly representing the recursive set of being an induction axiom of PA.

LEMMA 5.7. $\mathsf{PT}^- \vdash \forall x\,(\mathsf{C_t}(x) \wedge \mathsf{IndAx_{pa}}(x) \rightarrow \mathsf{T}(x))$.

Proof. Since $\mathsf{IndAx_{pa}}(x)$ is a representation of the induction axioms we have that PT^- proves:

$$(+)\ \forall x\,(\mathsf{IndAx_{pa}}(x) \rightarrow \exists y\,(\mathsf{Fml}^1(y) \wedge x = [y(\mathsf{0}) \dot{\wedge} \dot{\forall} z\,(y(\dot{z}) \dot{\rightarrow} y(\dot{\mathsf{S}}\dot{z})) \dot{\rightarrow} \dot{\forall} z\,y(\dot{z})])).$$

We argue in PT^-. Assuming that $\mathsf{C_t}(x) \wedge \mathsf{IndAx_{pa}}(x)$ we get that

$$\exists y\,(\mathsf{Fml}^1(y) \wedge x = [y(\mathsf{0}) \dot{\wedge} \dot{\forall} z\,(y(z) \dot{\rightarrow} y(\dot{\mathsf{S}}\dot{z})) \dot{\rightarrow} \dot{\forall} z\,y(z)]).$$

Moreover, by $\mathsf{pg}(y) \leq \mathsf{pg}(x)$ and $\mathsf{C_t}(x)$ we get $\mathsf{tc}(y)$. But then by $\mathsf{I_t I}$ also

$$\mathsf{T}(y(\mathsf{0})) \wedge \forall z\,(\mathsf{T}(y(\dot{z})) \rightarrow \mathsf{T}(y(\mathsf{num}(\mathsf{S}z)))) \rightarrow \forall z\,(\mathsf{T}(y(\dot{z}))).$$

By $\mathsf{tc}(y)$ and regularity and the fact that $\mathsf{PA} \vdash \mathsf{val}(\mathsf{num}(\mathsf{S}z)) = \mathsf{val}(\dot{\mathsf{S}}\dot{z})$ we have

$$\mathsf{T}(y(\mathsf{0})) \wedge \forall z\,(\mathsf{T}(y(\dot{z})) \rightarrow \mathsf{T}(y(\dot{\mathsf{S}}\dot{z}))) \rightarrow \forall z\,(\mathsf{T}(y(\dot{z}))).$$

[27] For a concrete version of a Hilbert style axiomatic calculus one can consider the system as presented in Pudlák (1998, p. 552f).

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

By $\mathsf{tc}(y)$ we have full compositionality for y and get

$$\mathsf{PT}^- \vdash \mathsf{T}[y(0) \mathbin{\dot\wedge} \dot\forall z\ (y(\dot z) \mathbin{\dot\to} y(\dot{\mathsf{S}}\,\dot z)) \mathbin{\dot\to} \dot\forall z\ y(\dot z)].$$

This suffices to get the desired result. □

It is easy to show that the truth for total formulas is closed under modus ponens.

LEMMA 5.8. $\mathsf{PT}^- \vdash \forall x \forall y\ (\mathsf{C_t}(x) \wedge \mathsf{C_t}(y) \wedge \mathsf{Tucl}(x) \wedge \mathsf{Tucl}(x \mathbin{\dot\to} y) \to \mathsf{Tucl}(y))$.

Also the rule of generalization restricted to the cut preserves truth.

LEMMA 5.9.

$$\mathsf{PT}^- \vdash \forall x \forall y\ (\mathsf{C_t}(x) \wedge \mathsf{C_t}(y) \wedge z \notin \mathsf{FV}(x) \wedge \mathsf{Tucl}(x \mathbin{\dot\to} y) \to \mathsf{Tucl}(x \mathbin{\dot\to} \dot\forall z\, y(z))).$$

Those lemmata will be sufficient to prove a restricted reflection principle.[28] Remember that $\mathsf{Pr_{pa}}(x, y)$ stands for there is a proof of y in PA of length less or equal to x.

THEOREM 5.10. *There is a cut* C *in* PT^- *such that* PT^- *proves*

$$\forall x \forall y\ (\mathsf{C}(x) \wedge \mathsf{C}(y) \wedge \mathsf{Pr_{pa}}(x, y) \to \mathsf{Tucl}(y)).$$

Proof. We will use the formula

$$\mathsf{I}(n) := \forall y\ (\mathsf{C_t}(y) \wedge \mathsf{Pr_{pa}}(n+1, y) \to \mathsf{Tucl}(y)).$$

First we will show that $\mathsf{I}(n)$ is inductive. $\mathsf{I}(0)$ is proved by Lemmas 5.6 and 5.7. For the induction step we use Lemmas 5.8 and 5.9.

By shortening we get a cut $\mathsf{C_I}(n)$ in PT^-, such that $\mathsf{C_I}(n) \to \mathsf{I}(n)$ and since it is a cut we get

$$\mathsf{C_I}(n) \to \forall y\ (\mathsf{C_t}(y) \wedge \mathsf{Pr_{pa}}(n+1, y) \to \mathsf{Tucl}(y)).$$

Then we get

$$\forall x \forall y\ (\mathsf{C_I}(x) \wedge \mathsf{C_t}(y) \wedge \mathsf{Pr_{pa}}(x, y) \to \mathsf{Tucl}(y)).$$

and by taking $\mathsf{C}(x) := \mathsf{C_I}(x) \wedge \mathsf{C_t}(x)$ we have our desired result. □

COROLLARY 5.11. *There is a cut* $\mathsf{C}(x)$ *in* PT^-*, such that*

$$\mathsf{PT}^- \vdash \forall x\ (\mathsf{C}(x) \to \mathsf{Con_{pa}}(x)).$$

Proof. Consider $0 = 1$. Clearly $\mathsf{PT}^- \vdash \mathsf{C}(\ulcorner 0 = 1 \urcorner)$ as it is a standard sentence. Moreover $\mathsf{PT}^- \vdash \neg\mathsf{T}(\ulcorner 0 = 1 \urcorner)$. By Theorem 5.10 we get $\forall z\ (\mathsf{C}(z) \to \neg\mathsf{Pr_{pa}}(z, \ulcorner 0 = 1 \urcorner))$, which is $\forall x\ (\mathsf{C}(x) \to \mathsf{Con_{pa}}(x))$ □

Then we can apply Theorem 4.12 to get

THEOREM 5.12. PT^- *has nonelementary speed-up over* PA.

So far we have only considered typed theories of truth. Most type-free theories of truth prove the consistency of PA such that we have nonelementary speed-up by Corollary 4.5. However, there is a version of KF, which is conservative over PA. This is Cantini's type-free theory $\mathsf{KF_t}$.[29] PT^- is basically a typed version of $\mathsf{KF_t}$. Since $\mathsf{PT}^- \subseteq \mathsf{KF_t}$ the speed-up result in Theorem 5.12 also holds for $\mathsf{KF_t}$.

[28] The proof is an adapted version of Heck's proof for Theorem 5.3. in Heck (2009).

[29] Compare Cantini (1989).

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

COROLLARY 5.13. $\mathsf{KF_t}$ *has nonelementary speed-up over* PA.

It is an open question whether $\mathsf{KF_t}$ has nonelementary speed-up over PT^- with respect to $\mathcal{L}_\mathsf{A}$ formulas. This would be interesting as it would provide a comparison of the expressiveness of typed and type-free theories of truth as both theories are proof-theoretically equivalent.

§6. Classical compositional truth. The case of classical compositional truth is a bit more complicated, and as far as the author knows it is not resolved. However, there are some partial answers due to observations by Albert Visser we want to mention. By classical compositional truth we mean the system that we get from positive truth by adding a clause for the commutation of negation and the truth predicate, that is, the universal closure of

$$\text{(CC)} \qquad \mathsf{Snt}(x) \to (\mathsf{T}(\dot{\neg} x) \leftrightarrow \neg \mathsf{T}x).$$

We get at least two systems depending on the induction axioms.

$$\mathsf{CT}\restriction := \mathsf{PA} \cup \text{(C1)-(C12)} \cup \text{(CC)};$$
$$\mathsf{CT} := \mathsf{PA} \cup \text{(C1)-(C12)} \cup \text{(CC)} \cup \text{ induction for } \mathcal{L}_\mathsf{T}\text{-formulas}.$$

There are more elegant axiomatizations of the system CT and its variants, but for our modest purposes the axiomatization will be sufficient.[30]

The axiom (CC) marks the difference between positive and classical compositional truth. (CC) itself is not a theorem of PT^-; however, there is a cut such that PT^- proves (CC) restricted to this cut. A proof of this is implicit in the proof of Lemma 5.4.

It is well known that CT proves the global reflection principle for PA and therefore also the consistency of PA.[31]

PROPOSITION 6.1. CT *proves* $\mathsf{Con_{pa}}$.

By applying Corollary 4.5 we see that

COROLLARY 6.2. CT *has nonelementary speed-up over* PA.

As far as the author knows the question whether $\mathsf{CT}\restriction$ has nonelementary speed-up is open. What is known is that it is impossible to use the same method of cuts with the consistency statements to show that there is nonelementary speed-up. The reason for this is the interpretability of $\mathsf{CT}\restriction$ in PA.[32]

PROPOSITION 6.3. $\mathsf{CT}\restriction$ *is relatively interpretable in* PA.

Due to an observation by Albert Visser (unpublished), the interpretability implies that there is no speed-up with regard to Π^0_1 formulas.

PROPOSITION 6.4 (Visser). $\mathsf{CT}\restriction$ *has no nonelementary speed-up over* PA *for* Π^0_1*-formulas.*

Proof. (Sketch) Let ψ be a Π^0_1-formula, that is ψ is of the form $\forall x\, \chi$ where χ is Δ_0. Let σ be an interpretation of $\mathsf{CT}\restriction$ in PA. Assume that $\mathsf{CT}\restriction \vdash^n \psi$. Then by the properties

[30] Compare Halbach (2011, p. 65). The system $\mathsf{PA}^{\mathsf{FT}}$, which is $\mathsf{PA}+$ 'there is a full truth class' is a variant of $\mathsf{CT}\restriction$ formulated in a relational language. Compare Enayat & Visser (2013).

[31] For a proof see Halbach (2011, p. 104f).

[32] For a proof see Enayat & Visser (2013).

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

of a relative interpretation there is a polynomial p such that $\mathsf{PA} \vdash^{p(n)} \sigma(\psi)$, that is, for $\delta(x)$ being the relativization of our translation $\mathsf{PA} \vdash^{p(n)} \forall x\,(\delta(x) \to \sigma(\chi))$.

We have to prove that also $\mathsf{PA} \vdash^{p(n)} \forall x\,\chi$ for some polynomial p. We will basically follow the proof strategy to show that an interpretation in PA preserves the Π^0_1 formulas.[33] Let $\mathfrak{M}$ be a PA model. Then the interpretation σ allows us to define a model $\mathfrak{M}^\sigma$. Moreover we can define an isomorphism f from $\mathfrak{M}$ to a model $\mathfrak{M}^f$, which is an initial segment of $\mathfrak{M}^\sigma$. But $\mathfrak{M}^\sigma \models \psi$ and by downward persistence of Π^0_1 formulas $\mathfrak{M}^f \models \psi$ and also $\mathfrak{M} \models \psi$. Since $\mathfrak{M}$ was an arbitrary PA model $\mathsf{PA} \vdash \psi$. We have to check that the relevant steps can be formalized in such a way that we get a polynomial bound on the length of proof.

The complexity of $\sigma(\chi)$ is restricted by some finite $k \leq p(n)$. Then by Theorem 3.1 we get a partial satisfaction predicate $\mathsf{Sat}_k(x, y)$ in PA such that PA proves the uniform Tarski biconditionals in a polynomial way. With this we get a polynomial p such that

$$\mathsf{PA} \vdash^{p(n)} \forall x\,(\delta(x) \to \mathsf{Sat}_k(\ulcorner \sigma\,\chi \urcorner, \mathsf{num}(x))).$$

Let F be a p.r. function and f its representation defined such that $\mathrm{F}(\#\mathsf{0}) = \#\mathsf{0}^\sigma$ and $\mathrm{F}(\#\mathsf{S}x) = \#\mathsf{S}^\sigma \mathsf{f}x$. Then $\mathsf{PA} \vdash^{p(n)} \delta(\mathsf{f}(\mathsf{0}))$ and $\mathsf{PA} \vdash^{p(n)} \delta(\mathsf{f}(x)) \to \delta(\mathsf{f}(\mathsf{S}x))$ by δ being the relativization of our interpretation of PA in PA. And with full induction in PA we get $\mathsf{PA} \vdash^{p(n)} \forall x\,\delta(\mathsf{f}(x))$ and therefore

$$\mathsf{PA} \vdash^{p(n)} \forall x\,\mathsf{Sat}_k(\ulcorner \sigma\,\chi \urcorner, \mathsf{num}(\mathsf{f}(x))).$$

Since f represents in a sense our isomorphism and χ being Δ_0 we also have[34]

$$\mathsf{PA} \vdash^{p(n)} \sigma\,\chi(\mathsf{f}(x)) \leftrightarrow \chi(x)).$$

and with the T-biconditionals also

$$\mathsf{PA} \vdash^{p(n)} \mathsf{Sat}_k(\ulcorner \sigma\,\chi \urcorner, \mathsf{num}(\mathsf{f}(x))) \to \mathsf{Sat}_k(\ulcorner \chi \urcorner, \mathsf{num}(x)).$$

This gives us

$$\mathsf{PA} \vdash^{p(n)} \forall x\,\mathsf{Sat}_k(\ulcorner \chi \urcorner, \mathsf{num}(x)).$$

and so we get our desired

$$\mathsf{PA} \vdash^{p(n)} \forall x\,\chi(x).$$

□

COROLLARY 6.5. $\mathsf{CT}\restriction$ *does not prove the consistency of* PA *on a cut.*

Another fact that is known about $\mathsf{CT}\restriction$ is an upper bound for speed-up.

PROPOSITION 6.6 (Leigh).[35] $\mathsf{CT}\restriction$ *attains at best superexponential speed-up.*

Although for $\mathsf{CT}\restriction$ the usual strategy doesn't work, there is a system for which it works. By adding the statement 'All axioms of PA are true.' to the axioms of $\mathsf{CT}\restriction$ we get a theory that is still a conservative extension of PA that has nonelementary speed-up over PA.

The proof of Proposition 6.1 contains at least two applications of induction for formulas containing the truth predicate. The first is needed to prove that 'All axioms of PA are true' and the second is needed to prove that for all n the result of a PA-proof of length

[33] Compare Hájek & Pudlák (1993, p. 170f).
[34] Compare Hájek & Pudlák (1993, p. 171).
[35] See for Leigh (2013, p. 17, Corollary 4).

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

n is true. $\mathsf{CT}\restriction$ can do neither of these proof steps. So adding the statement 'All axioms of PA are true.' to $\mathsf{CT}\restriction$ really makes a difference, although we still stay conservative over PA. 'All axioms of PA are true' is the statement $\forall x\,(\mathsf{Snt}(x) \wedge \mathsf{Ax}_{\mathsf{pa}}(x) \rightarrow \mathsf{T}(x))$. So let CT^* be $\mathsf{CT}\restriction + \forall x\,(\mathsf{Snt}(x) \wedge \mathsf{Ax}_{\mathsf{pa}}(x) \rightarrow \mathsf{T}(x))$. The method of Kotlarski *et al.* (1981) of extending recursively saturated models can be used to show the following.

PROPOSITION 6.7. CT^* *is a conservative extension of* PA.

On the other hand CT^* is able to prove the consistency of PA on a cut.

PROPOSITION 6.8. CT^* *proves the consistency of* PA *on a cut.*[36]

From this and Corollary 6.5 it follows that $\mathsf{CT}\restriction$ does not prove 'All axioms of PA are true' and by Theorem 4.12 we have:

COROLLARY 6.9. CT^* *has nonelementary speed-up over* PA.

§7. Philosophical implications. Let us clarify the relevance of nonelementary speed-up as a criterion for axiomatic theories truth in the light of the technical observations. For theories of truth that are able to prove $\mathsf{Con}_{\mathsf{pa}}$, nonelementary speed-up as a criterion is not too distinctive, because all those theories will satisfy it. The case of axiomatic theories of truth that are conservative extensions of PA is more interesting. One of the motivations for considering conservative extensions might be a deflationist conception of truth. So even if the technical results are interesting for primitivist accounts that are in favor of axiomatic theories but assign no special role to the truth predicate, the criterion becomes especially relevant for deflationist accounts. Deflationists claim that the utility of the truth predicate consists in its expressive function and moreover that the expressive function is the only function it has, it is so to say its *raison d'être*. Whereas the shortening of proofs is clearly useful, the connection between speed-up and expressivity is not as obvious.

However if we look at nonconservative extensions we see that the reason for speed-up is basically the possibility of proving the generalization $\forall x\,\mathsf{Con}_{\mathsf{pa}}(x)$ and with it the instances $\mathsf{Con}_{\mathsf{pa}}(\underline{n})$ for arbitrarily large n by a proof polynomial in n. Here we see the connection between speed-up and generalizations. Analysing the speed-up results for conservative extensions shows that although we are not able to prove $\forall x\,\mathsf{Con}_{\mathsf{pa}}(x)$ with an unrestricted universal quantifier, we can prove it for an universal quantifier restricted to a cut, that is, we can prove a generalization, which in this case is restricted to a proper definable cut. Not all the cuts definable by the truth predicate are proper and it is only possible to have a proper definable cut if the induction is restricted to a proper part of the language. As our base theory PA has induction for the language of arithmetic we need a language expansion to enable a proper definable cut. The truth predicate allows us to define this proper cut, which we were not able to define in the language of arithmetic. The introduction of the truth predicate indicates therefore an increase of expressiveness and this fits perfectly with the deflationist claims.

In general we therefore take nonelementary speed-up to be a positive feature of a theory of truth, as it partly explains the utility of the truth predicate. So we should add it to our preferred list of desiderata, such as for example the list of desiderata discussed in Leitgeb (2007). Especially for deflationist theories it should be added as a relevant criterion.

[36] Heck (2009) has a proof of generalization of this theorem in a draft for base theories containing Q. He gives credit for the result to Visser.

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

Accepting nonelementary speed-up as a criterion for deflationist theories allows us to draw two conclusions. On the one hand our results indicate that KF_t, PT^- and CT^* are promising candidates for deflationist theories of truth. On the other hand the lack of significant speed-up is a sign of an inadequacy of the proposed axiomatization. Polynomial speed-up is not significant and is not sufficient for a sign of increase of expressive power. Our results in case of disquotational theories such as TB and UTB therefore point towards an inadequacy of those theories regarding the expressive function of truth. This finding is in line with common criticism of disquotationalism.[37]

In the remainder we want to address some possible objections. One might question the relevance of the speed-up results because the theorems for which we get speed-up are not 'natural' arithmetical statements, but only syntactical consistency statements. However this form of criticism is not compelling in the case of truth. Here the consistency statements are especially important and intimately connected with the notion of truth. So the speed-up of truth theories for consistency statements is particularly relevant.[38] Moreover it is not clear that the speed-up phenomenon is restricted to syntactical statements. In the cases considered, the class of sentences for which we get speed-up contained partial consistency statements. These are restricted forms of $\mathsf{Con}_{\mathsf{pa}}$, a sentence that is independent from PA and the independence plays a crucial role in Friedman and Pudlák's proof that there are no short proofs of the partial consistency statements in PA.[39] It seems possible that there are other independent arithmetical statements, which maybe more 'natural' and allow for partial versions, such that we also get speed-up results for the class of those arithmetical statements.

Another line of criticism might question the robustness of the results. In this paper, the speed-up phenomenon is only investigated with respect to the base theory PA. One question is whether we get similar results for other base theories. We think that this is a question that needs addressing, but at the moment we can only conjecture that the speed-up results also hold for set theoretic base theories that are reflexive. In the case of weaker base theories with restricted induction like $\mathsf{I}\Sigma_n$ we think that the proofs are not directly transferable. One reason is that without the reflexivity of the base theory not all the partial consistency statements will be provable in the base theory. Another reason is more specific to theories of truth with the induction axiom $\mathsf{I}_t\mathsf{I}$, as for example PT^-. For PT^- it makes no difference if we start with PA as base theory or $\mathsf{I}\Sigma_n$, we always get all induction axioms for the arithmetical language.[40] So because of the reflexivity of PA we will be able to prove the consistency of the base theory and therefore our theory of truth will not be conservative over the base theory. This observation seems to support the claim that in the case of truth theories, PA is a natural base theory. Base theories with restricted forms of induction such as $\mathsf{I}\Sigma_n$ might appear artificial in the context of truth theories, because the truth predicate ignores the quantifier complexities of arithmetical sentences in the sense that independent of the complexity of an arithmetical sentence φ, the provably equivalent sentence $\mathsf{T}(\ulcorner\varphi\urcorner)$ is atomic. The complexity of the arithmetical sentence is so to say hidden in the term

37 Compare Gupta (1993) and Halbach (1999, 2011, chap. 7).

38 Thanks to Sean Walsh for the helpful suggestion.

39 The role of independent statements for speed-up is for example expressed by Pudlák, who claims that 'all concrete examples of speed-up have been constructed from concrete independent sentences by instantiation.' (Pudlák, 2013, p. 499)

40 Compare Heck (2009).

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

we apply the truth predicate to. This becomes apparent if we look at the induction axiom $\mathsf{I}_\mathsf{t}\mathsf{I}$ of PT^-.

As a third problem we want to comment on a general objection of deflationism itself. There is a famous argument that tries to establish the inadequacy of deflationism in connection with the conservativity claim. There are different versions of this conservativeness argument in the literature, going back to Shapiro (1998) and Ketland (1999). According to the argument, deflationist theories are on the one hand committed to conservative extensions of PA and on the other hand to theories that adequately express the soundness of PA. The arguments usually identify adequately expressing the soundness of PA and proving global reflection for PA, that is $\forall x\,(\mathsf{Snt}(x) \wedge \mathsf{Pr}_{\mathsf{pa}}(x) \rightarrow \mathsf{T}(x))$. To satisfy both requirements is impossible for theories of truth that contain the T-biconditionals for arithmetical sentences. However this identification is not necessary. There might be other ways to adequately express the soundness of PA. In the course of proving the consistency of PA on a cut some conservative truth theories were able to prove a global reflection principle on a cut, that is $\forall x \forall y\,(\mathsf{Snt}(x) \wedge \mathsf{C}(x) \wedge \mathsf{C}(y) \wedge \mathsf{Pr}_{\mathsf{pa}}(x, y) \rightarrow \mathsf{T}(x))$. Now it could be argued that global reflection on a cut is a generalization that adequately expresses a soundness claim for the base theory. Although this new line of response is an interesting byproduct of our investigation it cannot be developed here at length.

A fourth worry is also specific to the deflationist conception. According to deflationism the truth predicate has no other function besides the expressive function. If speed-up is not only part of the expressive function of truth but a sign of substance, then many theories of truth must be deemed unacceptable by deflationist standards. Although the talk about 'substance' is not sufficiently clear, it is hard to see why speed-up would make truth substantial. A possible line of reasoning for the substantialist claim would be to understand a simplification of proofs to be explanatory. Shortening of proofs is possibly not the only kind of simplification of proofs. There might be simplifications that give insights into why the theorem holds. A possible simplification of this kind could be a conceptual simplification.[41] For such a simplification an explanatory understanding might be justified. However, in the case of speed-up itself it is doubtful whether such an explanatory aspect is at work. And even if one is convinced that speed-up is explanatory this should not be a serious worry. Speed-up itself is not a mysterious property that would inflate truth and should therefore be acceptable by deflationist standards. One could even go a step further and give deflationism an instrumental twist. On this conception the usefulness of the truth predicate would be emphasized. The truth predicate is only introduced for instrumental purposes, such as shortening proofs. However, this expansion with the truth predicate should be unproblematic and beyond doubt of consistency. In this case it would be justified to replace the mysterious nonsubstantiality claim by the commitment to conservative extensions. This enables a philosophical position that avoids a deflationist discussion whether speed-up is merely expressive or also has an explanatory function.

§8. Acknowledgments. Research on this article was partly funded by the DFG within the project 'Syntactical Treatments of Interacting Modalities' and by the Alexander von Humboldt Foundation. A previous version of this paper was presented at the *Axiomatic Theories of Truth* conference in Oxford 2011. The author is grateful to the audience of the conference and the anonymous referees for helpful comments. Thanks to Albert Visser for

[41] Compare Caldon & Ignjatović (2005) for a discussion of conceptual simplification.

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

the permission to mention his unpublished observations and also to Richard Heck for the permission to cite his unpublished draft.

BIBLIOGRAPHY

Boolos, G. S. (1984). Don't eliminate cut. *Journal of Philosophical Logic*, **13**, 373–378.

Burgess, J. P. (2005). *Fixing Frege*. Princeton, NJ: Princeton University Press.

Buss, S. R. (1994). On Gödel's theorems on length of proofs I: Number of lines and speedup for arithmetics. *The Journal of Symbolic Logic*, **59**(3), 737–756.

Caldon, P., & Ignjatović, A. (2005). On mathematical instrumentalism. *Journal of Symbolic Logic*, **70**(3), 778–794.

Cantini, A. (1989). Notes on formal theories of truth. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, **35**, 97–130.

Enayat, A., & Visser, A. (2014). New constructions of satisfaction classes. In Achourioti, T., Galinon, H., Fujimoto, K., and Martínez-Fernández, J., editors. *Unifying the Philosophy of Truth.* Berlin: Springer.

Fujimoto, K. (2010). Relative truth definability of axiomatic truth theories. *Bulletin of Symbolic Logic*, **16**(3), 305–344.

Gödel, K. (1936). Über die Länge von Beweisen. In *Ergebnisse eines mathematischen Kolloquiums*, pp. 23–24. Reprinted in Feferman, S., Dawson, J., Kleene, S., Moore, G., Solovay, R., and Heijenoort, J., editors. *Kurt Gödel: Collected Works*, Vol. I. Publications 1929–1936. Oxford: Oxford University Press.

Gupta, A. (1993). A critique of deflationism. *Philosophical Topics*, **21**(1), 57–81.

Hájek, P., & Pudlák, P. (1993). *Metamathematics of First-Order Arithmetic*. Berlin: Springer Verlag.

Halbach, V. (1999). Disquotationalism and infinite conjunctions. *Mind*, **108**, 1–22.

Halbach, V. (2011). *Axiomatic Theories of Truth*. Cambridge, UK: Cambridge University Press.

Heck, R. G. (2009). The strength of truth-theories. draft.

Horsten, L. (1995). The semantical paradoxes, the neutrality of truth, and the neutrality of the minimalist theory of truth. In Cortois, P., editor. *The Many Problems of Realism*, Tilburg: Tilburg University Press, pp. 173–187.

Horsten, L. (2011). *The Tarskian Turn. Deflationism and Axiomatic Truth*. Cambridge, MA: MIT Press.

Horwich, P. (1998). *Truth* (second edition). Oxford, UK: Clarendon Press.

Ketland, J. (1999). Deflationism and Tarski's paradise. *Mind*, **108**, 69–94.

Kotlarski, H., Krajewski, S., & Lachlan, A. (1981). Construction of satisfaction classes for non-standard models. *Canadian Mathematical Bulletin*, **24**, 283–293.

Leigh, G. E. (2013). Conservativity for theories of compositional truth via cut elimination. arXiv:1308.0168 [math.LO].

Leitgeb, H. (2007). What theories of truth should be like (but cannot be). *Philosophy Compass*, **2**(2), 276–290.

McGee, V. (1992). Maximal consistent sets of instances of Tarski's schema (T). *The Journal of Philosophical Logic*, **21**, 235–241.

Paris, J., & Dimitracopoulos, C. (1983). A note on the undefinability of cuts. *The Journal of Symbolic Logic*, **48**(3), 564–569.

Pudlák, P. (1985). Cuts, consistency statements and interpretations. *The Journal of Symbolic Logic*, **50**, 423–441.

Pudlák, P. (1998). The lengths of proofs. In Buss, S. R., editor. *Handbook of Proof Theory*, Chapter VIII, Amsterdam, Netherlands: Elsevier Science Publisher. pp. 547–637.

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press

Pudlák, P. (2013). *Logical Foundations of Mathematics and Computational Complexity*. London: Springer.
Quine, W. V. O. (1970). *Philosophy of Logic*. Cambridge, MA: Harvard University Press.
Shapiro, S. (1998). Proof and truth: Through thick and thin. *The Journal of Philosophy*, **95**, 493–521.
Tarski, A., Mostowski, A., & Robinson, R. (1953). *Undecidable Theories*. Amsterdam, Netherlands: North-Holland Publishing Company.

MUNICH CENTER FOR MATHEMATICAL PHILOSOPHY
LUDWIG-MAXIMILIANS-UNIVERSITY
80539 MUNICH, GERMANY
E-mail: M.Fischer@lrz.uni-muenchen.de

https://doi.org/10.1017/S1755020314000070 Published online by Cambridge University Press