The Bulletin of Symbolic Logic
Volume 26, Number 3-4, September-December 2020

FINDING THE LIMIT OF INCOMPLETENESS I

YONG CHENG

Abstract. In this paper, we examine the limit of applicability of Gödel's first incompleteness theorem (G1 for short). We first define the notion "G1 holds for the theory T". This paper is motivated by the following question: can we find a theory with a minimal degree of interpretation for which G1 holds. To approach this question, we first examine the following question: is there a theory T such that Robinson's **R** interprets T but T does not interpret **R** (i.e., T is weaker than **R** w.r.t. interpretation) and G1 holds for T? In this paper, we show that there are many such theories based on Jeřábek's work using some model theory. We prove that for each recursively inseparable pair $\langle A, B\rangle$, we can construct a r.e. theory $U_{\langle A,B\rangle}$ such that $U_{\langle A,B\rangle}$ is weaker than **R** w.r.t. interpretation and G1 holds for $U_{\langle A,B\rangle}$. As a corollary, we answer a question from Albert Visser. Moreover, we prove that for any Turing degree $\mathbf{0} < \mathbf{d} < \mathbf{0}'$, there is a theory T with Turing degree $\mathbf{d}$ such that G1 holds for T and T is weaker than **R** w.r.t. Turing reducibility. As a corollary, based on Shoenfield's work using some recursion theory, we show that there is no theory with a minimal degree of Turing reducibility for which G1 holds.

§1. Introduction. Gödel's incompleteness theorem is one of the most remarkable results in the foundation of mathematics and has had great influence in logic, philosophy, mathematics, physics, and computer science, as discussed in [25] and [31]. Gödel proved his incompleteness theorems in [12] for a certain formal system **P** related to Russell–Whitehead's Principia Mathematica and based on the simple theory of types over the natural number series and the Dedekind–Peano axioms (see [1], p. 3). The following theorem is a modern reformulation of Gödel's first incompleteness theorem (where **PA** refers to the first-order theory commonly known as Peano Arithmetic).

Theorem 1.1 (Gödel–Rosser, first incompleteness theorem (G1).) *If T is a recursively axiomatized consistent extension of* **PA**, *then T is incomplete.*

The following is a well known open question about G1.

Question 1.2. Exactly how much arithmetical information from **PA** is needed for the proof of G1?

The notion of interpretation provides us a method to compare different theories in different languages (for the definition of interpretation, see

Received October 29, 2019.
2020 Mathematics Subject Classification. 03F40, 03F30, 03D35.
Key words and phrases. Gödel's first incompleteness theorem, interpretation, essential undecidability, Robinson's **R**.

1079-8986/20/2603-4-0007
DOI:10.1017/bsl.2020.9

Section 2). Given theories S and T, $S \unlhd T$ denotes that S is interpretable in T (or T interprets S), and $S \lhd T$ denotes that S is interpretable in T but T is not interpretable in S. We say that the theory S and T are *mutually interpretable* if $S \unlhd T$ and $T \unlhd S$. In this paper, we equate a set of sentences Γ in the language of arithmetic with the set of Gödel's numbers of sentences in Γ (see Section 2 for more details about Gödel's number). Given two arithmetic theories U and V, $U \leq_T V$ denotes that the theory U as a set of natural numbers is Turing reducible to the theory V as a set of natural numbers, and $U <_T V$ denotes that $U \leq_T V$ but $V \nleq_T U$.

Note that G1 can be generalized via interpretability: there exists a weak recursively axiomatizable consistent subtheory T (e.g., Robinson Arithmetic **Q**) of **PA** such that for each recursively axiomatizable consistent theory S, if T is interpretable in S, then S is incomplete (see [35]). To generalize this fact, in the following, we propose a general new notion "G1 holds for the theory T".

DEFINITION 1.3.

(1) Let T be a recursively axiomatizable consistent theory. We say that G1 holds for T if for any recursively axiomatizable consistent theory S, if T is interpretable in S, then S is incomplete.
(2) We say the theory S has a *minimal degree of interpretation* if there is no theory T such that $T \lhd S$.
(3) We say the theory S has a *minimal degree of Turing reducibility* if there is no theory V such that $V <_T S$.
(4) In this paper, whenever we say that *the theory S is weaker than the theory T w.r.t. interpretation* , this means that $S \lhd T$.
(5) In this paper, whenever we say that *the theory S is weaker than the theory V w.r.t. Turing reducibility* , this means that $S <_T V$.

Toward Question 1.2, in this project, we want to examine the following question:

QUESTION 1.4. Can we find a theory S such that G1 holds for S and S has a minimal degree of interpretation?

It is well known that G1 holds for Robinson Arithmetic **Q** (see [35]). From [35], G1 also holds for Robinson's theory **R** (for definitions of **Q** and **R**, we refer to Section 2). In Section 2, we review some theories mutually interpretable with **Q** for which G1 holds, and some theories mutually interpretable with **R** for which G1 holds. As the first step toward Question 1.4, we propose the following question:

QUESTION 1.5. Can we find a theory S such that G1 holds for S and $S \lhd \mathbf{R}$?

We find that Jeřábek essentially answered this question in [21]. In this paper, we show that there are many examples of such a theory S: for each recursively inseparable pair $\langle A, B\rangle$, we can construct a r.e. theory $U_{\langle A,B\rangle}$ such that G1 holds for $U_{\langle A,B\rangle}$ and $U_{\langle A,B\rangle} \lhd \mathbf{R}$ based on Jeřábek's work using some model theory. As a corollary, we answer a question from Albert Visser. For

Question 1.4, if we consider the degree of Turing reducibility instead of the degree of interpretation, the answer becomes easier. We show that for any Turing degree $\mathbf{0} < \mathbf{d} < \mathbf{0}'$, there is a theory S such that G1 holds for S, $S <_T \mathbf{R}$ and S has Turing degree $\mathbf{d}$ based on Shoenfield's work using some recursion theory. As a corollary, there is no theory with a minimal degree of Turing reducibility for which G1 holds.

The structure of this paper is as follows. In Section 1, we introduce our research questions and main results of this paper. In Section 2, we list some basic notions and facts we use in this paper, and give a review of theories weaker than **PA** w.r.t. interpretation from the literature for which G1 holds. In Section 3, we examine the limit of applicability of G1 w.r.t. interpretation. We prove Theorem 3.12, and answer a question from Albert Visser. In Section 4, we examine the limit of applicability of G1 w.r.t. Turing reducibility, and prove Theorem 4.5 and Corollary 4.6.

§2. Preliminaries. In this section, we review some basic notions and facts used in this paper. Our notations are standard. For books on Gödel's incompleteness theorem, we refer to [8], [25], [24], [31], [2] and [4]. For survey papers on Gödel's incompleteness theorem, we refer to [1], [23], [32], [41], [5] and [6]. For meta-mathematics of subsystems of **PA**, we refer to [18].

In this paper, a language consists of an arbitrary number of relation and function symbols of arbitrary finite arity.[1] For a given theory T, we use $L(T)$ to denote the language of T, and often equate $L(T)$ with the list of nonlogical symbols of the language. For a formula ϕ in $L(T)$, let $T \vdash \phi$ denote that ϕ is provable in T (i.e., there is a finite sequence of formulas $\langle \phi_0, \ldots, \phi_n \rangle$ such that $\phi_n = \phi$, and for any $0 \leq i \leq n$, either ϕ_i is an axiom of T or ϕ_i follows from some ϕ_j $(j < i)$ by using one inference rule). Theory T is *consistent* if no contradiction is provable in T. A formula ϕ is *independent* of T if $T \nvdash \phi$ and $T \nvdash \neg\phi$. A theory T is *incomplete* if there is a sentence ϕ in $L(T)$ such that ϕ is independent of T; otherwise, T is *complete* (i.e., for any sentence ϕ in $L(T)$, either $T \vdash \phi$ or $T \vdash \neg\phi$).

In this paper, we understand that each theory T comes with a fixed arithmetization. Let T be a recursively axiomatizable theory. Under this fixed arithmetization, we could establish the one-to-one correspondence between formulas of $L(T)$ and natural numbers. Under this correspondence, we can translate metamathematical statements about the formal theory T into statements about natural numbers. Furthermore, fundamental metamathematical relations can be translated in this way into certain recursive relations, hence into relations representable in the theory T. Consequently, one can speak about a formal system of arithmetic and about its properties as a theory in the system itself (see [25])! This is the essence of Gödel's idea of arithmetization. Under arithmetization, any formula or finite sequence of formulas can be coded by a natural number (this code

[1] We may view nullary functions as constants and nullary relations as propositional variables.

https://doi.org/10.1017/bsl.2020.9 Published online by Cambridge University Press

is called a Gödel number). In this paper, we use $\ulcorner\phi\urcorner$ to denote the Gödel number of ϕ. For details of arithmetization, we refer to [25].

Given a set of sentences Σ, we say Σ is *recursive* if the set of Gödel numbers of sentences in Σ is recursive. A theory T is *decidable* if the set of sentences provable in T is recursive; otherwise it is *undecidable*. A theory T is *recursively axiomatizable* if it has a recursive set of axioms (i.e., the set of Gödel numbers of axioms of T is recursive), and it is *finitely axiomatized* if it has a finite set of axioms. A theory T is *recursively enumerable* (r.e.) if it has a recursively enumerable set of axioms. A theory T is *essentially undecidable* if any recursively axiomatizable consistent extension of T in the same language is undecidable. A theory T is *essentially incomplete* if any recursively axiomatizable consistent extension of T in the same language is incomplete. The theory of completeness/incompleteness is closely related to the theory of decidability/undecidability. A theory T is *minimal essentially undecidable* if T is essentially undecidable, and if deleting any axiom of T, the remaining theory is no longer essentially undecidable. A theory T is *locally finitely satisfiable* if every finitely axiomatized subtheory of T has a finite model.

A n-ary relation $R(x_1,\dots,x_n)$ on $\mathbb{N}^n$ is *representable* in T iff there is a formula $\phi(x_1,\dots,x_n)$ such that $T\vdash\phi(\overline{m_1},\dots,\overline{m_n})$ if $R(m_1,\dots,m_n)$ holds (for $n\in\mathbb{N}$, we denote by $\overline{n}$ the corresponding numeral for n in $L(\mathbf{PA})$), and $T\vdash\neg\phi(\overline{m_1},\dots,\overline{m_n})$ if $R(m_1,\dots,m_n)$ does not hold. We say that a partial function $f(x_1,\dots,x_n)$ on $\mathbb{N}^n$ is representable in T iff there is a formula $\varphi(x_1,\dots,x_n,y)$ such that $T\vdash\forall y(\varphi(\overline{a_1},\dots,\overline{a_n},y)\leftrightarrow y=\overline{m})$ whenever $a_1,\dots,a_n,m\in\mathbb{N}$ are such that $f(a_1,\dots,a_n)=m$.

For the definitions of translation and interpretation, we follow [21]. Let T be a theory in a language $L(T)$, and S a theory in a language $L(S)$. In its most simple form, a *translation* I of language $L(T)$ into language $L(S)$ is specified by:

(1) an $L(S)$-formula $\delta_I(x)$ denoting the domain of I;
(2) for each relation symbol R of $L(T)$, an $L(S)$-formula R_I of the same arity;
(3) for each function symbol F of $L(T)$ of arity k, an $L(S)$-formula F_I of arity $k+1$.

If ϕ is an $L(T)$-formula, its I-translation ϕ^I is an $L(S)$-formula constructed as follows: we reformulate the formula in an equivalent way so that function symbols only occur in atomic subformulas of the form $F(\overline{x})=y$ where x_i,y are variables; then we replace each such atomic formula with $F_I(\overline{x},y)$, we replace each atomic formula of the form $R(\overline{x})$ with $R_I(\overline{x})$, and we restrict all quantifiers and free variables to objects satisfying δ_I. Moreover, we rename bound variables to avoid variable clashes during the process (see [21]).

A translation I of $L(T)$ into $L(S)$ is an *interpretation* of T in S if S proves:

https://doi.org/10.1017/bsl.2020.9 Published online by Cambridge University Press

(1) for each function symbol F of $L(T)$ of arity k, the formula expressing that F_I is total on δ_I: $\forall x_0, \dots \forall x_{k-1}(\delta_I(x_0) \wedge \cdots \wedge \delta_I(x_{k-1}) \rightarrow \exists y(\delta_I(y) \wedge F_I(x_0, \dots, x_{k-1}, y)))$;
(2) the I-translations of all theorems of T, and axioms of equality.

The simplified picture of translations and interpretations above actually describes only *one-dimensional*, *parameter-free*, and *one-piece* translations (see [21]). For the precise and technical definitions of a *multidimensional* interpretation, an interpretation *with parameters*, and a *piece-wise* interpretation, we refer to [40], [38] and [39] for the details.

A theory T is *interpretable* in a theory S if there exists an interpretation of T in S. If T is interpretable in S, then all sentences provable (refutable) in T are mapped, by the interpretation function, to sentences provable (refutable) in S. Interpretability can be accepted as a measure of strength of different theories. If $S \lhd T$, then S can be considered weaker than T w.r.t. interpretation; if S and T are mutually interpretable, then T and S are equally strong w.r.t. interpretation. The theory U *weakly interprets* the theory V (or V is weakly interpretable in U) if V is interpretable in some consistent extension of U in the same language (or equivalently, for some interpretation τ, the theory $U + V^\tau$ is consistent).

A general method for establishing the undecidability of theories is developed in [35]. The following theorem provides us two methods to prove the essentially undecidability of a theory via interpretation and representability.

THEOREM 2.1.

(1) ([35, Theorem 7, p. 22]) *Let T_1 and T_2 be two theories such that T_1 is consistent, and T_2 is interpretable in T_1. We then have: if T_2 is essentially undecidable, then T_1 is also essentially undecidable.*
(2) ([35, Corollary 2, p. 49]) *If T is a consistent theory in which all recursive functions are representable, then T is essentially undecidable.*

In Section 3, we will show that G1 holds for the theory T iff T is essentially undecidable. In the following, we review some theories from the literature which are weaker than **PA** w.r.t. interpretation, and which are essentially undecidable (i.e., G1 holds for them).

Robinson Arithmetic **Q** was introduced in [35] by Tarski et al. as a base axiomatic theory for investigating incompleteness and undecidability.

DEFINITION 2.2. Robinson Arithmetic **Q** is defined in the language $\{\mathbf{0}, \mathbf{S}, +, \cdot\}$ with the following axioms:

$\mathbf{Q}_1$: $\forall x \forall y(\mathbf{S}x = \mathbf{S}y \rightarrow x = y)$;
$\mathbf{Q}_2$: $\forall x(\mathbf{S}x \neq \mathbf{0})$;
$\mathbf{Q}_3$: $\forall x(x \neq \mathbf{0} \rightarrow \exists y(x = \mathbf{S}y))$;
$\mathbf{Q}_4$: $\forall x(x + \mathbf{0} = x)$;
$\mathbf{Q}_5$: $\forall x \forall y(x + \mathbf{S}y = \mathbf{S}(x + y))$;
$\mathbf{Q}_6$: $\forall x(x \cdot \mathbf{0} = \mathbf{0})$;
$\mathbf{Q}_7$: $\forall x \forall y(x \cdot \mathbf{S}y = x \cdot y + x)$.

https://doi.org/10.1017/bsl.2020.9 Published online by Cambridge University Press

Robinson Arithmetic **Q** is very weak and inadequate to formalize arithmetic: for instance, **Q** does not even prove that addition is associative. Robinson showed that any consistent theory that interprets **Q** is undecidable, and hence **Q** is essentially undecidable (see [35]). The fact that **Q** is essentially undecidable is very useful, and can be used to prove the essentially undecidability of other theories via Theorem 2.1. Since **Q** is finitely axiomatized, it follows that any theory that weakly interprets **Q** is undecidable. In fact, **Q** is minimal essentially undecidable in the sense that if deleting any axiom of **Q**, then the remaining theory is not essentially undecidable, and has a complete decidable extension (see [35, Theorem 11, p. 62]). Nelson [26] embarked on a program of investigating how much mathematics can be interpreted in Robinson Arithmetic **Q**: what can be interpreted in **Q**, but also what cannot be interpreted in **Q**. In fact, **Q** represents a rich degree of interpretability since a lot of stronger theories are interpretable in it as it can be shown (e.g., using Solovay's method of shortening cuts (see [15]), one can show that **Q** interprets fairly strong theories like $I\Delta_0+\Omega_1$ on a definable cut). The Lindenbaum algebras of all recursively enumerable theories that interpret **Q** are recursively isomorphic (see Pour-El and Kripke [27]).

The theory **PA** consists of axioms $\mathbf{Q}_1$–$\mathbf{Q}_2$, $\mathbf{Q}_4$–$\mathbf{Q}_7$ in Definition 2.2 and the following axiom scheme of induction: $(\phi(\mathbf{0})\wedge\forall x(\phi(x)\rightarrow\phi(\mathbf{S}x)))\rightarrow\forall x\,\phi(x)$, where ϕ is a formula with at least one free variable x.

Now we first discuss some prominent fragments of **PA** extending **Q** from the literature.

We define the arithmetic hierarchy $I\Sigma_n$ and $B\Sigma_n$ in the language of **PA**. An $L(\mathbf{PA})$-formula is bounded (or Δ_0 formula) if all quantifiers occurring in it are bounded, that is, in the form $(\exists x\leq y)\phi$ and $(\forall x\leq y)\phi$. For the definitions of Σ_n,Π_n, and Δ_n formulas $(n\geq 1)$, we refer to [25]. *Collection* for Σ_{n+1} formulas is the following principle:

$$(\forall x<u)(\exists y)\varphi(x,y)\rightarrow(\exists v)(\forall x<u)(\exists y<v)\varphi(x,y),$$

where $\varphi(x,y)$ is a Σ_{n+1} formula possibly containing parameters distinct from u,v.

The theory $I\Sigma_n$ is **Q** plus induction for Σ_n formulas and $B\Sigma_{n+1}$ is $I\Sigma_0$ plus collection for Σ_{n+1} formulas. It is well known that the following theories form a strictly increasing hierarchy:

$$I\Sigma_0,\ B\Sigma_1,\ I\Sigma_1,\ B\Sigma_2,\ \ldots I\Sigma_n,\ B\Sigma_{n+1},\ldots,\mathbf{PA}.$$

Define $\omega_1(x)=x^{|x|}$ and $\omega_{n+1}(x)=2^{\omega_n(|x|)}$, where $|x|$ is the length of the binary expression of x. Note that the graphs of these functions can be defined in our language with the recursive defining equation provable (see [18]). Let Ω_n denote the statement $\forall x\exists y(\omega_n(x)=y)$ which says that $\omega_n(x)$ is total. There is a bounded formula $\mathsf{Exp}(x, y, z)$ such that $I\Sigma_0$ proves that $\mathsf{Exp}(x, 0, z)\leftrightarrow z=1$ and $\mathsf{Exp}(x, Sy, z)\leftrightarrow\exists t(\mathsf{Exp}(x, y, t)\wedge z=t\cdot x)$ (see [10, Proposition 2, p. 299]). However, $I\Sigma_0$ cannot prove the totality of $\mathsf{Exp}(x, y, z)$. Let **exp** denote the statement postulating the totality of the exponential function $\forall x\forall y\exists z\,\mathsf{Exp}(x, y, z)$.

https://doi.org/10.1017/bsl.2020.9 Published online by Cambridge University Press

THEOREM 2.3 ([17], [10].)

(1) *For any* $n \geq 1$, $I\Sigma_0 + \Omega_n$ *is interpretable in* $\mathbf{Q}$ (*see* [10, Theorem 3, p. 304]).
(2) $I\Sigma_0 + \mathbf{exp}$ *is not interpretable in* $\mathbf{Q}$.[2]
(3) $I\Sigma_1$ *is not interpretable in* $I\Sigma_0 + \mathbf{exp}$ (*see* [17, Theorem 1.1, p. 186]).
(4) $I\Sigma_{n+1}$ *is not interpretable in* $B\Sigma_{n+1}$ (*see* [17, Theorem 1.2, p. 186]).
(5) $B\Sigma_1 + \mathbf{exp}$ *is interpretable in* $I\Sigma_0 + \mathbf{exp}$ (*see* [17, Theorem 2.4, p. 188]).
(6) *For each* $n \geq 1$, $B\Sigma_1 + \Omega_n$ *is interpretable in* $I\Sigma_0 + \Omega_n$ (*see* [17, Theorem 2.5, p. 189]).
(7) *For each* $n \geq 0$, $B\Sigma_{n+1}$ *is interpretable in* $I\Sigma_n$ (*see* [17, Theorem 2.6, p. 189]).

As a corollary, we have:

(1) The theories $\mathbf{Q}, I\Sigma_0, I\Sigma_0 + \Omega_1, \ldots, I\Sigma_0 + \Omega_n, \ldots, B\Sigma_1, B\Sigma_1 + \Omega_1, \ldots, B\Sigma_1 + \Omega_n, \ldots$ are all mutually interpretable;
(2) $I\Sigma_0 + \mathbf{exp}$ and $B\Sigma_1 + \mathbf{exp}$ are mutually interpretable;
(3) For $n \geq 1$, $I\Sigma_n$ and $B\Sigma_{n+1}$ are mutually interpretable;
(4) $\mathbf{Q} \lhd I\Sigma_0 + \mathbf{exp} \lhd I\Sigma_1 \lhd I\Sigma_2 \lhd \cdots \lhd I\Sigma_n \lhd \cdots \lhd \mathbf{PA}$.

Now we discuss some weak theories from the literature which are mutually interpretable with $\mathbf{Q}$.

It is interesting to compare $\mathbf{Q}$ with its bigger brother $\mathbf{PA}^-$. The theory $\mathbf{PA}^-$ is the theory of commutative, discretely ordered semirings with a minimal element plus the subtraction axiom. The theory $\mathbf{PA}^-$ has the following axioms with $L(\mathbf{PA}^-) = L(\mathbf{PA}) \cup \{\leq\}$: (1) $x + 0 = x$; (2) $x + y = y + x$; (3) $(x + y) + z = x + (y + z)$; (4) $x \cdot 1 = x$; (5) $x \cdot y = y \cdot x$; (6) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$; (7) $x \cdot (y + z) = x \cdot y + x \cdot z$; (8) $x \leq y \vee y \leq x$; (9) $(x \leq y \wedge y \leq z) \rightarrow x \leq z$; (10) $x + 1 \nleq x$; (11) $x \leq y \rightarrow (x = y \vee x + 1 \leq y)$; (12) $x \leq y \rightarrow x + z \leq y + z$; (13) $x \leq y \rightarrow x \cdot z \leq y \cdot z$; (14) $x \leq y \rightarrow \exists z(x + z = y)$. From [40], $\mathbf{PA}^-$ is interpretable in $\mathbf{Q}$, and hence $\mathbf{PA}^-$ is mutually interpretable with $\mathbf{Q}$.

Let $\mathbf{Q}^+$ be the extension of $\mathbf{Q}$ with the following extra axioms ($L(\mathbf{Q}^+) = L(\mathbf{Q}) \cup \{\leq\}$):

$\mathbf{Q}_8$: $(x + y) + z = x + (y + z)$;
$\mathbf{Q}_9$: $x \cdot (y + z) = x \cdot y + x \cdot z$;
$\mathbf{Q}_{10}$: $(x \cdot y) \cdot z = x \cdot (y \cdot z)$;
$\mathbf{Q}_{11}$: $x + y = y + x$;
$\mathbf{Q}_{12}$: $x \cdot y = y \cdot x$;
$\mathbf{Q}_{13}$: $x \leq y \leftrightarrow \exists z(x + z = y)$.

The theory $\mathbf{Q}^+$ is interpretable in $\mathbf{Q}$ (see [10, Theorem 1, p. 296]), and hence $\mathbf{Q}^+$ is mutually interpretable with $\mathbf{Q}$.

Andrzej Grzegorczyk considered a theory $\mathbf{Q}^-$ in which addition and multiplication do satisfy natural reformulations of axioms of $\mathbf{Q}$ but are possibly nontotal functions. More exactly, the language of $\mathbf{Q}^-$ is $\{\mathbf{0}, \mathbf{S}, A, M\}$

[2] See [10, Theorem 6, p. 313]. Solovay proved that $I\Sigma_0 + \neg\mathbf{exp}$ is interpretable in $\mathbf{Q}$ (see [10, Theorem 7, p. 314]).

where A and M are ternary relations, and the axioms of $\mathbf{Q}^-$ are the axioms $\mathbf{Q}_1$– $\mathbf{Q}_3$ of **Q** plus the following six axioms about A and M:

A: $\forall x \forall y \forall z_1 \forall z_2 (A(x,y,z_1) \wedge A(x,y,z_2) \rightarrow z_1 = z_2)$;
M: $\forall x \forall y \forall z_1 \forall z_2 (M(x,y,z_1) \wedge M(x,y,z_2) \rightarrow z_1 = z_2)$;
G4: $\forall x\, A(x,\mathbf{0},x)$;
G5: $\forall x \forall y \forall z (\exists u (A(x,y,u) \wedge z = \mathbf{S}(u)) \rightarrow A(x,\mathbf{S}(y),z))$;
G6: $\forall x\, M(x,\mathbf{0},\mathbf{0})$;
G7: $\forall x \forall y \forall z (\exists u (M(x,y,u) \wedge A(u,x,z)) \rightarrow M(x,\mathbf{S}(y),z))$.

Andrzej Grzegorczyk asked whether $\mathbf{Q}^-$ is essentially undecidable. Petr Hájek considered a somewhat stronger theory with axioms

H5: $\forall x \forall y \forall z (\exists u (A(x,y,u) \wedge z = \mathbf{S}(u)) \Leftrightarrow A(x,\mathbf{S}(y),z))$ and
H7: $\forall x \forall y \forall z (\exists u (M(x,y,u) \wedge A(u,x,z)) \leftrightarrow M(x,\mathbf{S}(y),z))$

instead of **G5** and **G7**. He showed that this stronger variant of $\mathbf{Q}^-$ is essentially undecidable (see [16]). Vítězslav Švejdar provided a positive answer to Grzegorczyk's original question in [33], and proved that **Q** is interpretable in $\mathbf{Q}^-$ using the Solovay's method of shortening cuts (and hence $\mathbf{Q}^-$ is essentially undecidable). Thus, $\mathbf{Q}^-$ is mutually interpretable with **Q**.

Andrzej Grzegorczyk proposed the theory of concatenation (**TC**) in [13] as a possible alternative theory for studying incompleteness and undecidability. Unlike Robinson (or Peano) Arithmetic, where the individuals are numbers that can be added or multiplied, in **TC** one has strings (or texts) that can be concatenated. We refer to [13] for Grzegorczyk's philosophical motivations to study **TC**.

The theory **TC** has the language $\{\frown, \alpha, \beta, =\}$ with a binary function symbol and two constants, and the following axioms:

TC1: $\forall x \forall y \forall z (x \frown (y \frown z) = (x \frown y) \frown z)$;
TC2: $\forall x \forall y \forall u \forall v (x \frown y = u \frown v \rightarrow ((x = u \wedge y = v) \vee \exists w ((u = x \frown w \wedge w \frown v = y) \vee (x = u \frown w \wedge w \frown y = v))))$;
TC3: $\forall x \forall y (\alpha \neq x \frown y)$;
TC4: $\forall x \forall y (\beta \neq x \frown y)$;
TC5: $\alpha \neq \beta$.

Grzegorczyk [13] proved (mere) undecidability of the theory **TC**. Grzegorczyk and Zdanowski [14] proved that **TC** is essentially undecidable. However, [14] leaves an interesting unanswered question: are **TC** and **Q** mutually interpretable? Švejdar [34] showed that $\mathbf{Q}^-$ is interpretable in **TC**, and hence **Q** is interpretable in **TC** since **Q** is interpretable in $\mathbf{Q}^-$. Ganea [11] gave a different proof of the interpretability of **Q** in **TC**, but he also used the detour via $\mathbf{Q}^-$. Visser [37] gave a proof of the interpretability of **Q** in **TC** not using $\mathbf{Q}^-$. Note that **TC** is easily interpretable in the bounded arithmetic $I\Sigma_0$. Thus, **TC** is mutually interpretable with **Q**.

Adjunctive Set Theory (**AS**) is the following theory in the language with only one binary relation symbol $\in$.

AS1: $\exists x \forall y (y \notin x)$;
AS2: $\forall x \forall y \exists z \forall u (u \in z \leftrightarrow (u = x \vee u = y))$.

The theory **AS** interprets Robinson's Arithmetic **Q**, and hence is essentially undecidable. Nelson [26] showed that **AS** is interpretable in **Q**. Thus, **AS** is mutually interpretable with **Q**.

The theory $\mathbf{S^1_2}$ is a finitely axiomatizable weak arithmetic introduced by Buss in [3] to study polynomial time computability. The theory $\mathbf{S^1_2}$ gives us what we need to formalize the proof of the second incompleteness theorem in a natural and effortless way. In fact, it is easier to do it in $\mathbf{S^1_2}$ than in **PA**, since the restrictions present in $\mathbf{S^1_2}$ prevent one from making wrong turns and inefficient choices (see [40]). From [10], $I\Sigma_0$ is interpretable in $\mathbf{S^1_2}$, and $\mathbf{S^1_2}$ is interpretable in **Q**. Thus, $\mathbf{S^1_2}$ is mutually interpretable with **Q**.

Now, we introduce Robinson's theory **R** introduced by Tarski et al. in [35], and some variants of **R**.

Definition 2.4. Let **R** be the theory consisting of schemes Ax1– Ax5 with $L(\mathbf{R}) = \{\mathbf{0}, \mathbf{S}, +, \cdot, \leq\}$ where $\leq$ is a primitive binary relation symbol, and $\overline{n} = \mathbf{S}^n\mathbf{0}$ for $n \in \mathbb{N}$:

Ax1: $\overline{m} + \overline{n} = \overline{m+n}$;
Ax2: $\overline{m} \cdot \overline{n} = \overline{m \cdot n}$;
Ax3: $\overline{m} \neq \overline{n}$, if $m \neq n$;
Ax4: $\forall x(x \leq \overline{n} \to x = \overline{0} \vee \cdots \vee x = \overline{n})$;
Ax5: $\forall x(x \leq \overline{n} \vee \overline{n} \leq x)$.

The axiom schemes of **R** contain all key properties of arithmetic for the proof of G1. The theory **R** is not finitely axiomatizable. Note that $\mathbf{R} \lhd \mathbf{Q}$ since **Q** is not interpretable in **R**: if **Q** is interpretable in **R**, then it is interpretable in some finite fragment of **R**; however, **R** is locally finitely satisfiable but any model of **Q** is infinite. Visser [39] proved the following universal property of **R** which provides a unique characterization of **R**.

Theorem 2.5 (Visser, Theorem 6, [39].) *For any r.e. theory T, T is locally finitely satisfiable iff T is interpretable in* **R**.[3]

We say a specific class Φ of sentences has the finite model property if every satisfiable sentence of Φ has a finite model. Since relational Σ_2 sentences in a finite relational language have the finite model property (see Chapter 5 in [7]), by Theorem 2.5, any consistent theory axiomatized by a recursive set of Σ_2 sentences in a finite relational language is interpretable in **R**. Since all recursive functions are representable in **R** (see [35, Theorem 6, p. 56]), from Theorem 2.1(2), **R** is essentially undecidable. Cobham showed that **R** has a stronger property than essentially undecidability.

Theorem 2.6 (Cobham, [36].) *Any r.e. theory that weakly interprets* **R** *is undecidable.*[4]

Now, we discuss some variants of **R**. If not explicitly mentioned otherwise, we assume that the base language is the same as $L(\mathbf{R}) = \{\mathbf{0}, \mathbf{S}, +, \cdot, \leq\}$ with

[3]In fact, if T is locally finitely satisfiable, then T is interpretable in **R** via a one-piece one-dimensional parameter-free interpretation.

[4]Vaught [36] gave a proof of Cobham's theorem via existential interpretation.

https://doi.org/10.1017/bsl.2020.9 Published online by Cambridge University Press

$\leq$ as a primitive binary relation symbol. Let $\mathbf{R}_0$ be the theory consisting of schemes Ax1, Ax2, Ax3, and Ax4. The theory $\mathbf{R}_0$ is no longer essentially undecidable: the theory $\mathbf{R}_0$ has a decidable complete extension given by the theory of real closed fields with $\leq$ as the empty relation on reals. In fact, whether $\mathbf{R}_0$ is essentially undecidable depends on the language of $\mathbf{R}_0$. If $L(\mathbf{R}_0) = \{\mathbf{0}, \mathbf{S}, +, \cdot\}$ with $\leq$ defined in terms of $+$, then $\mathbf{R}_0$ is essentially undecidable: Cobham first observed that $\mathbf{R}$ is interpretable in $\mathbf{R}_0$, and hence $\mathbf{R}_0$ is essentially undecidable (see [36] and [22]). Let $\mathbf{R}_1$ be the theory consisting of schemes Ax1, Ax2, Ax3, and Ax4′, where Ax4′ is defined as follows:

$$\forall x(x \leq \overline{n} \leftrightarrow x = \overline{0} \vee \cdots \vee x = \overline{n}).$$

The theory $\mathbf{R}_1$ is essentially undecidable since $\mathbf{R}$ is interpretable in $\mathbf{R}_1$ (see [22], p. 62). However, $\mathbf{R}_1$ is not minimal essentially undecidable. Let $\mathbf{R}_2$ be the system consisting of schemes Ax2, Ax3, and Ax4′. From [22], $\mathbf{R}$ is interpretable in $\mathbf{R}_2$, and hence $\mathbf{R}_2$ is essentially undecidable.[5] The theory $\mathbf{R}_2$ is minimal essentially undecidable in the sense that if we delete any axiom scheme of $\mathbf{R}_2$, then the remaining theory is not essentially undecidable: if we delete Ax2, then the theory of natural numbers with $x \cdot y$ defined as $x + y$ is a complete decidable extension; if we delete Ax3, then the theory of models with only one element is a complete decidable extension; if we delete Ax4′, then the theory of real closed fields is a complete decidable extension. By essentially the same argument as the proof of Theorem 2.6 in [40], we can show that any r.e. theory that weakly interprets $\mathbf{R}_2$ is undecidable.

Kojiro Higuchi and Yoshihiro Horihata introduced the theory of concatenation $\mathbf{WTC}^{-\varepsilon}$, which is a weak subtheory of Grzegorczyk's theory $\mathbf{TC}$, and showed that $\mathbf{WTC}^{-\varepsilon}$ is minimal essentially undecidable and $\mathbf{WTC}^{-\varepsilon}$ is mutually interpretable with $\mathbf{R}$ (see [19]).

Elementary Arithmetic ($\mathbf{EA}$) is $I\Delta_0 + \mathbf{exp}$. We refer to [18] for the definition of Primitive Recursive Arithmetic ($\mathbf{PRA}$). In summary, we have the following pictures.

(1) Theories $\mathbf{PA}^-, \mathbf{Q}^+, \mathbf{Q}, \mathbf{Q}^-, \mathbf{TC}, \mathbf{AS}$, and $\mathbf{S}^1_2$ are mutually interpretable and are all essentially undecidable.
(2) Theories $\mathbf{R}, \mathbf{R}_1, \mathbf{R}_2$, and $\mathbf{WTC}^{-\varepsilon}$ are mutually interpretable and are all essentially undecidable.
(3) $\mathbf{R} \lhd \mathbf{Q} \lhd \mathbf{EA} \lhd \mathbf{PRA} \lhd \mathbf{PA}$.

Now, a natural question is: among finitely axiomatized theories for which G1 holds, does $\mathbf{Q}$ have the least degree of interpretation? The following theorem tells us that the answer is no.

THEOREM 2.7 (Visser, Theorem 2, [40].) *Suppose $\mathbf{R} \subseteq A$, where A is finitely axiomatized and consistent. Then, there is a finitely axiomatized B such that $\mathbf{R} \subseteq B \subseteq A$ and $B \lhd A$.*

[5] Another way to show that $\mathbf{R}_2$ is essentially undecidable is to prove that all recursive functions are representable in $\mathbf{R}_2$.

Theorem 2.7 shows that the structure $\langle\{S : \mathbf{R} \unlhd S \lhd \mathbf{Q}\}, \lhd\rangle$ is not well founded w.r.t. finitely axiomatized theories.

THEOREM 2.8 (Visser, Theorem 12, [40].) *Suppose A and B are finitely axiomatized theories that interpret $\mathbf{S^1_2}$. Then there are finitely axiomatized theories $\overline{A} \supseteq A$ and $\overline{B} \supseteq B$ such that $\overline{A}$ and $\overline{B}$ are incomparable (i.e., $\overline{A} \ntrianglelefteq \overline{B}$ and $\overline{B} \ntrianglelefteq \overline{A}$).*

Theorem 2.8 shows that there are many pairs of incomparable theories extending $\mathbf{Q}$.

§3. Finding the limit of applicability of G1 w.r.t. interpretation. In this section, we examine the limit of applicability of G1 w.r.t. interpretation, and show that we can find many theories weaker than $\mathbf{R}$ w.r.t. interpretation for which G1 holds based on Jeřábek's work using some model theory. First of all, we give some equivalent characterizations of the notion "G1 holds for the theory T".

PROPOSITION 3.1. *Let T be a recursively axiomatizable consistent theory. The following are equivalent:*

(1) *G1 holds for T.*
(2) *T is essentially incomplete.*
(3) *T is essentially undecidable.*

PROOF. (1) $\Rightarrow$ (2) is trivial.

(2) $\Leftrightarrow$ (3): It is well known that every consistent recursively axiomatizable complete theory is decidable; and every incomplete decidable theory has a consistent, decidable complete extension in the same language (see Corollary 3.1.8 and Theorem 3.1.9 in [25], pp. 214–215). From these two facts, T is essentially undecidable iff T is essentially incomplete.

(2) $\Rightarrow$ (1): Follows from Theorem 2.1 and (2) $\Leftrightarrow$ (3). ⊣

As a corollary of Section 2, we have:

(1) G1 holds for the following theories, and they are mutually interpretable: $\mathbf{Q}, I\Sigma_0, I\Sigma_0 + \Omega_n (n \geq 1), B\Sigma_1, B\Sigma_1 + \Omega_n (n \geq 1), \mathbf{TC}, \mathbf{Q}^-, \mathbf{Q}^+, \mathbf{PA}^-, \mathbf{S^1_2}$.
(2) G1 holds for the following theories, and they are mutually interpretable: $\mathbf{R}, \mathbf{R}_1, \mathbf{R}_2$, and $\mathbf{WTC}^{-\varepsilon}$.

Up to now, we do not have any example of an essentially undecidable theory S such that $S \lhd \mathbf{R}$ and G1 holds for S. We find that Jeřábek's work in [21] essentially provides such an example of theory S. The motivating question of [21] is: if a theory represents all partial recursive functions, does it interpret Robinson's theory $\mathbf{R}$? Jeřábek [21] negatively answered this question, and showed that there exists a theory T in which all partial recursive functions are representable, yet T does not interpret $\mathbf{R}$. Jeřábek's proof uses tools from model theory: investigating model-theoretic properties of the model completion of the empty theory in a language with function symbols (see [21]).[6] Now we introduce Jeřábek's theory $\mathbf{Rep}_{\mathsf{PRF}}$. Let PRF denote the set of all partial recursive functions. The language $L(\mathbf{Rep}_{\mathsf{PRF}})$ consists of constant symbols $\overline{n}$ for each $n \in \mathbb{N}$, and function symbols $\overline{f}$ of appropriate arity for each partial recursive function f. The theory $\mathbf{Rep}_{\mathsf{PRF}}$ is axiomatized by:

(1) $\overline{n} \neq \overline{m}$ for $n \neq m \in \mathbb{N}$;
(2) $\overline{f}(\overline{n_0}, \ldots, \overline{n_{k-1}}) = \overline{m}$ for each k-ary partial recursive function f such that $f(n_0, \ldots, n_{k-1}) = m$, where $n_0, \ldots, n_{k-1}, m \in \mathbb{N}$.

Corollary 3.2. *G1 holds for* $\mathbf{Rep}_{PRF}$, *and* $\mathbf{Rep}_{PRF} \lhd \mathbf{R}$.

Proof. The theory $\mathbf{Rep}_{\mathsf{PRF}}$ is essentially undecidable since all recursive functions are representable in it. Since $\mathbf{Rep}_{\mathsf{PRF}}$ is locally finitely satisfiable, by Theorem 2.5, $\mathbf{Rep}_{\mathsf{PRF}} \unlhd \mathbf{R}$. Jeřábek [21] showed that $\mathbf{R}$ is not interpretable in $\mathbf{Rep}_{\mathsf{PRF}}$. Thus, G1 holds for $\mathbf{Rep}_{\mathsf{PRF}}$, and $\mathbf{Rep}_{\mathsf{PRF}} \lhd \mathbf{R}$. ⊣

Jeřábek's [21] is not written in the spirit of answering Question 1.5, and the potential of the method in [21] is not yet fully explored. Now, we give more examples of a theory S such that G1 holds for S and $S \lhd \mathbf{R}$ based on Jeřábek's work in [21].

Definition 3.3. We say $\langle S, T\rangle$ is a recursively inseparable pair if S and T are disjoint r.e. subsets of $\mathbb{N}$, and there is no recursive set $X \subseteq \mathbb{N}$ such that $S \subseteq X$ and $X \cap T = \emptyset$.

Definition 3.4. Let $\langle A, B\rangle$ be a recursively inseparable pair. Consider the following r.e. theory $U_{\langle A,B\rangle}$ with $L(U_{\langle A,B\rangle}) = \{\mathbf{0}, \mathbf{S}, \mathbf{P}\}$ where $\mathbf{P}$ is a unary relation symbol, and $\overline{n} = \mathbf{S}^n\mathbf{0}$ for $n \in \mathbb{N}$:

(1) $\overline{m} \neq \overline{n}$ if $m \neq n$;
(2) $\mathbf{P}(\overline{n})$ if $n \in A$;
(3) $\neg\mathbf{P}(\overline{n})$ if $n \in B$.

In the following, let $\langle A, B\rangle$ be an arbitrary recursively inseparable pair.

Lemma 3.5. *G1 holds for* $U_{\langle A,B\rangle}$.

Proof. By Proposition 3.1, it suffices to show that $U_{\langle A,B\rangle}$ is essentially incomplete. Let S be a recursively axiomatizable consistent extension of $U_{\langle A,B\rangle}$. Let $X = \{n : S \vdash P(\overline{n})\}$ and $Y = \{n : S \vdash \neg P(\overline{n})\}$. Then $A \subseteq X$ and $B \subseteq Y$. Since S is recursively axiomatizable and consistent, X and Y are disjoint recursive enumerable sets. Since $\langle A, B\rangle$ is recursively inseparable, $X \cup Y \neq \mathbb{N}$. Take $n \notin X \cup Y$. Then $S \nvdash P(\overline{n})$ and $S \nvdash \neg P(\overline{n})$. Hence S is incomplete. ⊣

Fact 3.6 (Theorem 2, p. 43, [24].). Let A and B be disjoint r.e. subsets of $\mathbb{N}$ and T be a consistent r.e. extension of $\mathbf{Q}$. Then there is a Σ_1 formula $\phi(x)$ such that for any n, we have:

[6] By the empty theory, we mean the theory with no extra logical axioms.

https://doi.org/10.1017/bsl.2020.9 Published online by Cambridge University Press

(1) $n \in A$ iff $T \vdash \phi(\overline{n})$;
(2) $n \in B$ iff $T \vdash \neg\phi(\overline{n})$.

LEMMA 3.7. *The theory $U_{\langle A,B\rangle}$ is interpretable in* **R**.

PROOF. By Fact 3.6, there exists a formula $\phi(x)$ with one free variable such that $\mathbf{R} \vdash \phi(\overline{n})$ iff $n \in A$, and $\mathbf{R} \vdash \neg\phi(\overline{n})$ iff $n \in B$.[7] Thus $U_{\langle A,B\rangle}$ is interpretable in **R** via interpreting $\mathbf{P}(x)$ as $\phi(x)$.

Here is another proof of Lemma 3.7: since $U_{\langle A,B\rangle}$ is a locally finitely satisfiable r.e. theory, by Theorem 2.5, $U_{\langle A,B\rangle}$ is interpretable in **R**. ⊣

Based on Jeřábek's work in [21], our strategy to prove that $U_{\langle A,B\rangle}$ does not interpret **R** is to consistently extend the interpreting theory to a theory with quantifier elimination, using the fact that the empty theory in an arbitrary language L has a model completion which we denote by $\mathbf{EC}_L$, the theory of existentially closed L-structures. The theory $\mathbf{EC}_L$ admits the elimination of quantifiers (see [21]). The two key tools we use to show that $U_{\langle A,B\rangle}$ does not interpret **R** are Theorem 3.8 and Theorem 3.9 which essentially use properties of $\mathbf{EC}_L$. We say a relation $R \subseteq X^2$ is *asymmetric* if there are no $a,b \in X$ such that $R(a,b)$ and $R(b,a)$.

THEOREM 3.8 (Jeřábek, Theorem 5.1, [21].) *For any first-order language L and formula $\phi(\overline{z},\overline{x},\overline{y})$ with $lh(\overline{x}) = lh(\overline{y})$, there is a constant n with the following property. Let $M \models EC_L$ and $\overline{u} \in M$ be such that $M \models \exists\overline{x}_0,\ldots,\exists\overline{x}_{n-1} \bigwedge_{i<j<n} \phi(\overline{u},\overline{x}_i,\overline{x}_j)$. Then for every $m \in \mathbb{N}$ and an asymmetric relation R on $\{0,\ldots,m-1\}$, $M \models \exists\overline{x}_0,\ldots,\exists\overline{x}_{m-1} \bigwedge_{\langle i,j\rangle \in R} \phi(\overline{u},\overline{x}_i,\overline{x}_j)$.*

THEOREM 3.9 (Jeřábek, Theorem 4.5, [21].) *For a Σ_2-axiomatized theory T, T is interpretable in a consistent existential theory iff T is weakly interpretable in* $\mathbf{EC}_L$ *for some language L.*

Especially, Jeřábek [21] showed that: (1) if a theory is interpretable in a consistent quantifier-free or existential theory, it is weakly interpretable in $\mathbf{EC}_L$ for some language L, and the interpretation can be taken quantifier-free; (2) if a Σ_2 theory is weakly interpretable in $\mathbf{EC}_L$, it is interpretable in a quantifier-free theory.

DEFINITION 3.10. Consider the following theory **T** in the language $\langle \in \rangle$ axiomatized by the sentences $\exists z, x_0,\ldots,x_{n-1}(\bigwedge_{i<j<n} x_i \neq x_j \bigwedge \forall y(y \in z \leftrightarrow \bigvee_{i<n} y = x_i))$ for all $n \in \mathbb{N}$.

PROPOSITION 3.11.

(1) **T** *is not weakly interpretable in* $\mathbf{EC}_L$ *for any language L.*
(2) **R** *is not weakly interpretable in* $\mathbf{EC}_L$ *for any language L.*
(3) *If* **R** *is interpretable in $U_{\langle A,B\rangle}$, then* **R** *is weakly interpretable in* $\mathbf{EC}_L$ *for some language L.*

[7]The proof of Fact 3.6 in [24, Theorem 2, p. 43] uses the fixed point theorem for the base theory T. Since the fixed point theorem holds for **R**, Fact 3.6 also applies to **R**.

PROOF. (1): Suppose this does not hold. Apply Theorem 3.8 to the formula which interprets $\bigwedge_{i<j<n} x_i \neq x_j \bigwedge \forall y(y \in z \leftrightarrow \bigvee_{i<n} y = x_i)$, and R a chain longer than n to get a contradiction.

(2): Note that $\mathbf{T}$ is interpretable in $\mathbf{R}$. Since $\mathbf{T}$ is not weakly interpretable in $\mathbf{EC}_L$ for any language L, $\mathbf{R}$ is not weakly interpretable in $\mathbf{EC}_L$ for any language L.

(3): This follows from Theorem 3.9 since $U_{\langle A,B\rangle}$ is a consistent r.e. theory. ⊣

THEOREM 3.12. *For any recursively inseparable pair $\langle A,B\rangle$, there is a r.e. theory $U_{\langle A,B\rangle}$ such that G1 holds for $U_{\langle A,B\rangle}$, and $U_{\langle A,B\rangle} \lhd \mathbf{R}$.*[8]

PROOF. By Proposition 3.11(2) and (3), $\mathbf{R}$ is not interpretable in $U_{\langle A,B\rangle}$. From Lemma 3.5 and Lemma 3.7, we have G1 holds for $U_{\langle A,B\rangle}$, and $U_{\langle A,B\rangle} \lhd \mathbf{R}$. ⊣

COROLLARY 3.13. *Let S be a consistent existential theory. Then the following are equivalent:*

(1) *G1 holds for S, and $S \lhd \mathbf{R}$ (i.e., S is a solution for Question 1.5);*
(2) *S is essentially undecidable and locally finitely satisfiable.*

PROOF. As a corollary of Theorem 3.9 and Proposition 3.11(2), any consistent existential theory does not interpret $\mathbf{R}$. Thus, from this, Proposition 3.1 and Theorem 2.5, we have the equivalence. ⊣

From Theorem 2.5, $\mathbf{R}$ has the universality property: every locally finitely satisfiable r.e. theory is interpretable in it. Albert Visser asked the following question:

QUESTION 3.14 (Visser). Would S with $S \unlhd \mathbf{R}$ such that G1 holds for S share the universality property of $\mathbf{R}$ that every locally finitely satisfiable r.e. theory is interpretable in it.

As a corollary of Theorem 3.12, the answer for this question is negative. We have shown that for any recursively inseparable pair $\langle A,B\rangle$, there is a theory $U_{\langle A,B\rangle}$ such that G1 holds for $U_{\langle A,B\rangle}$, and $U_{\langle A,B\rangle} \lhd \mathbf{R}$. The theory $\mathbf{R}$ is locally finitely satisfiable, but $\mathbf{R}$ is not interpretable in $U_{\langle A,B\rangle}$. Take another example: the theory $\mathbf{T}$ as in Definition 3.10 is locally finitely satisfiable, but $\mathbf{T}$ is not interpretable in $U_{\langle A,B\rangle}$ (if $\mathbf{T}$ is interpretable in $U_{\langle A,B\rangle}$, by Theorem 3.9, $\mathbf{T}$ is weakly interpretable in $\mathbf{EC}_L$ for some language L which contradicts Proposition 3.11(1)). Thus, for any recursively inseparable pair $\langle A,B\rangle$, the theory $U_{\langle A,B\rangle}$ is a counterexample for Visser's Question. This shows the speciality of $\mathbf{R}$: Theorem 2.5 provides a unique characterization of $\mathbf{R}$.

Define $\mathsf{D} = \{S : S \lhd \mathbf{R}$ and G1 holds for the theory $S\}$. We have shown that we could find many witnesses for D. We could naturally examine the

[8]However, Theorem 3.12 does not tell us more information about the theory $U_{\langle S,T\rangle}$ and $U_{\langle U,V\rangle}$ for different recursively inseparable pairs $\langle S,T\rangle$ and $\langle U,V\rangle$: for example, whether $U_{\langle S,T\rangle}$ and $U_{\langle U,V\rangle}$ have the same degree of interpretation or the same degree of Turing reducibility.

structure of $\langle \mathsf{D}, \lhd \rangle$. A natural question is: whether the similar results as in Theorem 2.7 and Theorem 2.8 apply to the structure $\langle \mathsf{D}, \lhd \rangle$. About the structure of $\langle \mathsf{D}, \lhd \rangle$, we could naturally ask:

QUESTION 3.15.

(1) Is $\langle \mathsf{D}, \lhd \rangle$ well founded (or is it that for any $S \in \mathsf{D}$, there is $T \in \mathsf{D}$ such that $T \lhd S$)?
(2) Are any two elements of $\langle \mathsf{D}, \lhd \rangle$ comparable (i.e., is it that for any $S, T \in \mathsf{D}$, we have either $S \unlhd T$ or $T \unlhd S$)?
(3) Could we find a theory S with a minimal degree of interpretation such that G1 holds for S?

In the rest of this paper, we will show that if we consider the Turing degree structure instead of the interpretation degree structure of D, we have definite answers for Question 3.15.

§4. The limit of applicability of G1 w.r.t. Turing reducibility. In this section, we examine the limit of applicability of G1 w.r.t. Turing reducibility, and show that there is no theory with a minimal degree of Turing reducibility for which G1 holds based on Shoenfield's work using some recursion theory.

Let $\mathcal{R}$ be the structure of the r.e. degrees with the ordering $\leq_T$ induced by Turing reducibility with the least element $\mathbf{0}$ and the greatest element $\mathbf{0}'$. Define $\overline{\mathsf{D}} = \{S : S <_T \mathbf{R}$, and G1 holds for the theory $S\}$. A natural question is to examine the structure of $\langle \overline{\mathsf{D}}, <_T \rangle$. Now, we will show that the structure $\langle \overline{\mathsf{D}}, <_T \rangle$ is much simpler than $\langle \mathsf{D}, \lhd \rangle$, and we have answers to Question 3.15 for the structure $\langle \overline{\mathsf{D}}, <_T \rangle$ based on Shoenfield's work using some recursion theory.

THEOREM 4.1 (Shoenfield, Theorem 1, [30].) *If A is recursively enumerable and not recursive, there is a recursively inseparable pair $\langle B, C \rangle$ such that A, B and C have the same Turing degree.*

Now, we will show that for any Turing degree $\mathbf{0} < \mathbf{d} < \mathbf{0}'$, there is a theory U such that G1 holds for U, $U <_T \mathbf{R}$, and U has Turing degree $\mathbf{d}$ (c.f. Theorem 4.5). The following theorem of Shoenfield is essential for the proof of Theorem 4.5. To make the reader have a better sense of how the theory U in Theorem 4.5 is constructed, we provide details of the proof of Theorem 4.2. Feferman [9] also proved that for any r.e. Turing degree one can design a formal theory whose corresponding decision problem is of the same degree (however, it is not clear whether such a formal theory is essentially undecidable).

THEOREM 4.2 (Shoenfield, Theorem 2, [30].) *Let A be recursively enumerable and not recursive. Then there is a consistent axiomatizable theory T having one nonlogical symbol which is essentially undecidable and has the same Turing degree as A.*

PROOF. By Theorem 4.1, pick a recursively inseparable pair $\langle B, C \rangle$ such that A, B, and C have the same Turing degree. Now we define the theory T

with $L(T) = \{R\}$, where R is a binary relation symbol. Theory T contains axioms asserting that R is an equivalence relation. Let Φ_n be the statement that there is an equivalence class of R consisting of n elements. Then, as axioms of T, we adopt Φ_n for all $n \in B$, and $\neg\Phi_n$ for all $n \in C$. Finally, for each n, we adopt an axiom asserting there is at most one equivalence class of R having n elements. Clearly, T is consistent and axiomatizable. Using models, we see Φ_n is provable iff $n \in B$, and $\neg\Phi_n$ is provable iff $n \in C$. Hence B and C are recursive in T.

Disjunctions of conjunctions whose terms are Φ_n or $\neg\Phi_n$ for some $n \in \mathbb{N}$, are called a disjunctive normal form of $\langle \Phi_n : n \in \mathbb{N} \rangle$.

LEMMA 4.3 (Janiczak, Lemma 2 in [20].) *Any sentence ϕ of the theory T is equivalent to a disjunctive normal form of $\langle \Phi_n : n \in \mathbb{N} \rangle$, and this disjunctive normal form can be found explicitly once ϕ is explicitly given.*[9]

By Lemma 4.3, every sentence ϕ of T is equivalent to a disjunctive normal form of $\langle \Phi_n : n \in \mathbb{N} \rangle$, and this disjunctive normal form can be calculated from ϕ. It follows that T is recursive in B and C. Hence T has the same Turing degree as A.

Finally, we show that T is essentially undecidable. Suppose T has a consistent decidable extension S. Let D be the set of n such that Φ_n is provable in S. Then D is recursive, $B \subseteq D$, and $C \cap D = \emptyset$ which contradicts the fact that $\langle B, C \rangle$ is a recursively inseparable pair. ⊣

THEOREM 4.4 (Sacks.)

(1) (*Embedding theorem,* [28]) *Every countable partial ordering can be embedded into $\mathcal{R}$.*
(2) (*Density theorem,* [29]) *For every pair of nonrecursive r.e. degrees $\mathbf{a} <_T \mathbf{b}$, there is one $\mathbf{c}$ such that $\mathbf{a} <_T \mathbf{c} <_T \mathbf{b}$.*

THEOREM 4.5. *For any Turing degree $\mathbf{0} < \mathbf{d} < \mathbf{0}'$, there is a theory U such that G1 holds for U, $U <_T \mathbf{R}$, and U has Turing degree $\mathbf{d}$.*

PROOF. From Theorem 4.2, for each Turing degree $\mathbf{0} < \mathbf{d} < \mathbf{0}'$, there is a theory U such that G1 holds for $U$and U has Turing degree $\mathbf{d}$. It is a well known fact that $\mathbf{R}$ has Turing degree $\mathbf{0}'$. ⊣

We could ask the similar question as Question 3.15 for the structure $\langle \overline{\mathsf{D}}, <_T \rangle$:

- Is $\langle \overline{\mathsf{D}}, <_T \rangle$ well founded?
- Are any two elements of $\langle \overline{\mathsf{D}}, <_T \rangle$ comparable?
- Could we find a theory S with a minimal degree of Turing reducibility such that G1 holds for S?

From Theorem 4.5 and Theorem 4.4, we have answers for these questions:

[9]This is a reformulation of Janiczak's Lemma 2 in [20] in the context of the theory T. Janiczak's Lemma is proved by means of a method known as the elimination of quantifiers.

https://doi.org/10.1017/bsl.2020.9 Published online by Cambridge University Press

COROLLARY 4.6.

(1) *The structure* $\langle \overline{\mathsf{D}}, <_T\rangle$ *is not well founded (i.e., for any* $S \in \overline{\mathsf{D}}$*, there is* $U \in \overline{\mathsf{D}}$ *such that* $U <_T S$*);*
(2) *The structure* $\langle \overline{\mathsf{D}}, <_T\rangle$ *has incomparable elements (i.e., there are* $U, V \in \overline{\mathsf{D}}$ *such that* $U \nleq_T V$ *and* $V \nleq_T U$*);*
(3) *There is no theory with a minimal degree of Turing reducibility for which G1 holds.*

In fact, we can improve Theorem 4.5 by making that the theory U is interpretable in $\mathbf{R}$.

THEOREM 4.7. *For any Turing degree* $\mathbf{0} < \mathbf{d} < \mathbf{0}'$*, there is a theory* U *such that G1 holds for* U*,* $U \trianglelefteq \mathbf{R}$ *and* U *has Turing degree* $\mathbf{d}$*.*

PROOF. Let $\mathbf{d}$ be a Turing degree with $\mathbf{0} < \mathbf{d} < \mathbf{0}'$. By Theorem 4.2, pick an essentially undecidable theory S with Turing degree $\mathbf{d}$.

Consider the product theory $S \otimes \mathbf{R}$ defined as follows. The theory $S \otimes \mathbf{R}$ has the following axioms: $P \rightarrow X$ if X is a S-axiom; $\neg P \rightarrow Y$ if Y is a $\mathbf{R}$-axiom, where P is a 0-ary predicate symbol.

Now, we show that $S \otimes \mathbf{R}$ is essentially undecidable (i.e., G1 holds for $S \otimes \mathbf{R}$) and interpretable in $\mathbf{R}$.

LEMMA 4.8. $S \otimes \mathbf{R}$ *is essentially undecidable.*

PROOF. Suppose U is a consistent decidable extension of $S \otimes \mathbf{R}$. Define $X = \{\langle \ulcorner \phi \urcorner, \ulcorner \psi \urcorner\rangle : U \vdash P \rightarrow \phi \text{ or } U \vdash \neg P \rightarrow \psi\}$. Since U is decidable, X is recursive. Note that $S \subseteq (X)_0$ and $\mathbf{R} \subseteq (X)_1$. We claim that at least one of $(X)_0$ and $(X)_1$ is consistent. If both $(X)_0$ and $(X)_1$ are inconsistent, then $U \vdash (P \rightarrow \perp)$ and $U \vdash (\neg P \rightarrow \perp)$. Thus, $U \vdash \perp$ which contradicts that U is consistent. WLOG, we assume that $(X)_0$ is consistent. Then $(X)_0$ is a consistent decidable extension of S, which contradicts that S is essentially undecidable. ⊣

It is easy to show that $S \otimes \mathbf{R}$ is interpretable in $\mathbf{R}$ (i.e., $S \otimes \mathbf{R} \trianglelefteq \mathbf{R}$): take the identity interpretation on the $\mathbf{R}$ side and interpret P as $\perp$.

Since S has Turing degree $\mathbf{d}$ and $\mathbf{R}$ has Turing degree $\mathbf{0}'$, $S \otimes \mathbf{R}$ has Turing degree $\mathbf{d}$. ⊣

However, from the proof of Theorem 4.7, we cannot get that $S \otimes \mathbf{R} \triangleleft \mathbf{R}$ (i.e., $\mathbf{R}$ is not interpretable in $S \otimes \mathbf{R}$). An interesting question is: could we improve Theorem 4.7, and show that for any Turing degree $\mathbf{0} < \mathbf{d} < \mathbf{0}'$, there is a theory U such that G1 holds for U, $U \triangleleft \mathbf{R}$ and U has Turing degree $\mathbf{d}$.

As far as we know, Question 3.15 is open. We make the conjecture that there is no theory with a minimal degree of interpretation for which G1 holds, $\langle \mathsf{D}, \triangleleft\rangle$ is not well founded and $\langle \mathsf{D}, \triangleleft\rangle$ has incomparable elements.

Acknowledgments. I would like to thank Prof. Albert Visser for inspiring discussions of the topic in this paper, insightful comments on the original version of the paper, and for introducing some papers to me. I would also like to thank Prof. Emil Jeřábek for his patient explanations of his work to me. I would like to thank Prof. Ali Enayat and Prof. V.Yu. Shavrukov for their comments on my paper. I would like to thank the referee for providing detailed and helpful comments for improvements. I thank the lamp for my feet, and the light on my path. This paper is the research result of The National Social Science Fund of China for general project "Research on the limit of incompleteness and the intensional problem of incompleteness" (project no: 18BZX131). I would like to thank the fund support by The National Social Science Fund of China for general project.

REFERENCES

[1] L. D. Beklemishev, *Gödel incompleteness theorems and the limits of their applicability I*. ***Russian Mathematical Surveys***, vol. 65 (2010), pp. 61–106.

[2] G. Boolos, ***The Logic of Provability***, Cambridge University Press, Cambridge, 1993.

[3] S. R. Buss, ***Bounded Arithmetic***, Bibliopolis, Napoli, 1986.

[4] Y. Cheng, ***Incompleteness for Higher-Order Arithmetic: An Example Based on Harrington's Principle***, Springer Series: Springerbrief in Mathematics, Springer, New York, 2019.

[5] ———, *Current research on Gödel's incompleteness theorems*, this Journal, to appear. arXiv:2009.04887v2

[6] ———, *On the depth of Gödel's incompleteness theorem*, this Journal, submitted. arXiv:2008.13142v1

[7] H.-D. Ebbinghaus and J. Flum, ***Finite Model Theory***, Springer Monographs in Mathematics, Springer, New York, 1999.

[8] H. B. Enderton, ***A Mathematical Introduction to Logic***, second ed., Academic Press, Boston, MA, 2001.

[9] S. Feferman, *Degrees of unsolvability associated with classes of formalized theories*. ***Journal of Symbolic Logic***, vol. 22 (1957), no. 2, pp. 161–175.

[10] F. Ferreira and G. Ferreira, *Interpretability in Robinson's* Q, this Journal, vol. 19 (2013), no. 3, pp. 289–317.

[11] M. Ganea, *Arithmetic on semigroups*. ***Journal of Symbolic Logic***, vol. 74 (2009), no. 1, pp. 265–278.

[12] K. Gödel, *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I*. ***Monatshefte für Mathematik***, vol. 38 (1931), no. 1, pp. 173–198.

[13] A. Grzegorczyk, *Undecidability without arithmetization*. ***Studia Logica***, vol. 79 (2005), no. 2, pp. 163–230.

[14] A. Grzegorczyk and K. Zdanowski, *Undecidability and concatenation*, ***Andrzej Mostowski and Foundational Studies*** (A. Ehrenfeucht, V. W. Marek, and M. Srebrny, editors), IOS Press, Amsterdam, 2008, pp. 72–91.

[15] D. Guaspari and R. M. Solovay, *Rosser sentences*. ***Annals of Mathematical Logic***, vol. 16 (1979), no. 1, pp. 81–99.

[16] P. Hájek, *Mathematical fuzzy logic and natural numbers*. ***Fundamenta Informaticae***, vol. 81 (2007), no. 1–3, pp. 155–163.

[17] ———, *Interpretability and fragments of arithmetic*, ***Arithmetic, Proof Theory and Computational Complexity*** (P. Clote and J. Jan Krajícek, editors). Oxford Logic Guides 23, Clarendon Press, Oxford, 1993, pp. 185–196.

[18] P. Hájek and P. Pudlák, ***Metamathematics of First-Order Arithmetic***, Springer-Verlag, Berlin-Heidelberg, 1993.

[19] K. Higuchi and Y. Horihata, *Weak theories of concatenation and minimal essentially undecidable theories-an encounter of WTC and S2S*. ***Archive for Mathematical Logic***, vol. 53 (2014), no. 7–8, pp. 835–853.

[20] A. Janiczak, *Undecidability of some simple formalized theories*. ***Fundamenta Mathematicae***, vol. 40 (1953), pp. 131–139.

[21] E. Jeřábek, *Recursive functions and existentially closed structures*. ***Journal of Mathematical Logic***, vol. 20 (2020), no. 1, 2050002.

[22] J. P. Jones and J. C. Shepherdson, *Variants of Robinson's essentially undecidable theory R*. ***Archive for Mathematical Logic***, vol. 23 (1983), pp. 65–77.

[23] H. Kotlarski, *The incompleteness theorems after 70 years*. ***Annals of Pure and Applied Logic,*** vol. 126 (2004), pp. 125–138.

[24] P. Lindström, ***Aspects of Incompleteness***, Lecture Notes in Logic, Springer-Verlag, Berlin, 1997.

[25] R. Murawski, ***Recursive Functions and Metamathematics: Problems of Completeness and Decidability, Gödel's Theorems***, Springer Netherlands, 1999.

[26] E. Nelson, ***Predicative Arithmetic***, Mathematical Notes, vol. 32, Princeton University Press, Princeton, NJ, 1986.

[27] M. B. Pour-El and S. Kripke, *Deduction-preserving "recursive isomorphisms" between theories*. ***Fundamenta Mathematicae***, vol. 61 (1967), pp. 141–163.

[28] G. E. Sacks, ***Degrees of Unsolvability,*** Annals of Mathematics Studies, vol. 55, Princeton University Press, Princeton, NJ, 1963.

[29] ———, *The recursively enumerable degrees are dense*. ***Annals of Mathematics***, vol. 80 (1964), no. 2, pp. 300–312.

[30] J. R. Shoenfield, *Degrees of formal systems*. ***The Journal of Symbolic Logic***, vol. 23 (1958), no. 4, pp. 389–392.

[31] P. Smith, ***An Introduction to Gödel's Theorems***, Cambridge University Press, Cambridge, 2007.

[32] C. Smoryński, *The incompleteness theorems*, ***Handbook of Mathematical Logic*** (J. Barwise, editor), North-Holland, Amsterdam, 1977, pp. 821–865.

[33] V. Švejdar, *An interpretation of Robinson arithmetic in its Grzegorczyk's weaker variant*. ***Fundamenta Informaticae***, vol. 81 (2007), no. 1–3, pp. 347–354.

[34] ———, *On interpretability in the theory of concatenation*. ***Notre Dame Journal of Formal Logic,*** vol. 50 (2009), no. 1, pp. 87–95.

[35] A. Tarski, A. Mostowski, and R. M. Robinson, ***Undecidabe Theories***, Studies in Logic and the Foundations of Mathematics, North-Holland, Amsterdam, 1953.

[36] R. L. Vaught, *On a theorem of Cobham concerning undecidable theories*, ***Proceedings of the 1960 International Congress on Logic, Methodology and Philosophy of Science*** (E. Nagel, P. Suppes, and A. Tarski, editors), Stanford University Press, Stanford, 1962, pp 14–25.

[37] A. Visser, *Growing commas: A study of sequentiality and concatenation*. ***Notre Dame Journal of Formal Logic***, vol. 50 (2009), no. 1, pp. 61–85.

[38] ———, *Can we make the second incompleteness theorem coordinate free?* ***Journal of Logic and Computation***, vol. 21 (2011), no. 4, pp. 543–560.

[39] ———, *Why the theory* ***R*** *is special?*, ***Foundational Adventures. Essay in Honour of Harvey Friedman***, College Publications, London, 2014, pp. 7–23.

[40] ———, *On Q*. ***Soft Computing,*** vol. 21 (2017), pp. 39–56.

[41] ———, *The second incompleteness theorem: Reflections and ruminations*, ***Gödel's Disjunction: The Scope and Limits of Mathematical Knowledge*** (L. Horsten and P. Welch, editors), Oxford University Press, Oxford, 2016.

SCHOOL OF PHILOSOPHY
WUHAN UNIVERSITY
WUHAN 430072, HUBEI
PEOPLE'S REPUBLIC OF CHINA
E-mail: world-cyr@hotmail.com