The Axiom System $I\Sigma_0$ Manages to Simultaneously Obey and Evade the Herbrandized Version of the Second Incompleteness Theorem

Dan E. Willard *

More polished version in *Electronic Notes in Theoretical Computer Science* 165 (November 2006) pp. 213–226 (This is the Proceedings of Wollics 2006.)

Abstract

In 1981, Paris and Wilkie [21] indicated it was an open question whether $I\Sigma_0$ would satisfy the Second Incompleteness Theorem for Herbrand deduction. We will show that $I\Sigma_0$ will both obey and defy the Herbrandized version of the Second Incompleteness Theorem, *depending on* which of several equivalent definitions of $I\Sigma_0$ one examines.

1 Introduction

Gödel's Second Incompleteness Theorem [9] asserts that neither Peano Arithmetic, nor any consistent extension of it, can prove a theorem affirming its own self-consistency under Hilbert deduction. There have been numerous generalizations and extensions of Gödel's seminal result [1, 2, 3, 4, 5, 6, 7, 8, 12, 15, 18, 19] [21, 22, 23, 24, 25, 26, 28, 27, 29, 30, 33, 35, 37, 39]. For example, the combined work of Pudlák and Solovay [23, 26] has shown that essentially no axiom system that recognizes $\text{Successor}(x) = x+1$ as a total function can prove a theorem affirming its own consistency under Hilbert deduction.

In 1981, Paris and Wilkie [21] noticed that it was an open question whether the axiom system $I\Sigma_0$ did satisfy the Second Incompleteness Theorem for cut-free methods of deduction. Interestingly, Paris-Wilkie observed that $I\Sigma_0$+Exp is unable to prove the Hilbert consistency of even an axiom system as simple as Q [31]. Subsequently. Adamowicz-Zbierski [1, 3] showed that $I\Sigma_0+\Omega_1$ was unable to verify its Herbrand and semantic tableaux consistency, and Willard [33, 35] expanded this result to show that the cut-free versions of the Second Incompleteness Theorem applied also to the standard textbook versions [10, 13, 17] of $I\Sigma_0$'s axiomatization.

On 16 November 2005, we received a fascinating email communication from L.A. Kołodziejczyk about this subject. It observed that there are two natural formalisms for axiomatizing $I\Sigma_0$,

*University of Albany, Email=dew@cs.albany.edu. Partially supported by NSF Grant CCR 99-02726

henceforth called Ax-1 and Ax-2. Both shall take the Tarski-Mostowski-Robinson axiom system Q as their starting base. In a context where $\phi(x,y)$ is a Δ_0 formula, these formalisms will use respectively Equations (1) and (2) to denote their induction schemes.

$$\forall x \quad \{ \quad \{ \ \phi(x,0) \ \ \wedge \ \ \forall y \, [\, \phi(x,y) \implies \phi(x,y') \ ] \ \} \quad \implies \quad \forall y \ \phi(x,y) \ \ \} \tag{1}$$

$$\forall x \, \forall z \, \{ \, \{ \, \phi(x,0) \ \wedge \ \forall y \le z \, [\, \phi(x,y) \implies \phi(x,y') \,] \, \} \implies \forall y \le z \, \phi(x,y) \, \} \tag{2}$$

Kołodziejczyk noticed that logically equivalent axiom systems, such as Ax-1 and Ax-2, do not necessarily have the same properties with regards to the cut-free versions of the Second Incompleteness Theorem. He thus asked whether [35]'s semantic tableaux version of the Second Incompleteness Theorem will generalize for Ax-2's unconventional induction scheme?

One half of our 2-part answer to this question appears in a separate paper [40]. It explains how our prior results about Ax-1's cut-free incompleteness properties have direct generalizations for Ax-2. The second half of our 2-part answer will appear in this conference paper. It is as follows:

Suppose one wishes to play the *very adversarial* role of being the Devil's Advocate who seeks to find other axiomatizations of $I\Sigma_0$, that prove the same theorems as Ax-1 and Ax-2. Then it turns out one can construct a third unorthodox axiomatization of $I\Sigma_0$, called Ax-3, which evades the Herbrandized version of the Second Incompleteness Theorem.

In order to understand the nature of this quite counter-intuitive effect, it is useful to recall that $\alpha \cong \beta$ denotes merely that the two axiom systems α and β formally prove the same set of theorems. The central point is that such an equivalence does not imply that these two systems can physically prove the statement that " $\alpha \cong \beta$ ". For this reason, one *certainly cannot automatically presume* that β satisfies a fixed version of the Second Incompleteness Theorem, *when* a logically equivalent axiom system α *does.* Thus, this paper will formalize a third equivalent axiomatization for $I\Sigma_0$ that manages to evade at least the Herbrandized version of the Second Incompleteness Theorem.

2 The Definition of a New Version of $I\Sigma_0$

A formula will be called Δ_0^R iff it has a structure similar to a Δ_0 formula except that its bounded quantifiers, " $\forall \ v \ \le \ T$ " and " $\exists \ v \ \le \ T,$ ", are now disallowed from using the conventional arithmetic functions of addition and multiplication in their terms T . Instead, the terms of a Δ_0^R formula will employ only the maximum function as the only permissible operator to define a variable's bounded range. (Arithmetic functions are allowed to appear elsewhere in the body of a Δ_0^R formula.) Thus, Equation (3) is an example of a Δ_0^R formula, and (4) is an example of a Δ_0 formula that is not Δ_0^R .

$$\forall \ p \le \ \text{Max}(x,y) \quad [\ (\ p+y \le \ x \ + \ 2*y) \ \vee \ (\ p*y \ \le \ y*y*y \ ) \ ] \tag{3}$$

$$\forall\, p \leq \; x * y \qquad [\; (\, p + y \leq \; x \, + \, 2 * y) \; \vee \; (\, p * y \; \leq \; y * y * x \;) \,] \tag{4}$$

Let us call a formula Π_1^R iff it can be written as $\forall\, v_1 \; \forall\, v_2 \; ... \; \forall\, v_n \;\; \phi(v_1, v_2, ...v_n)$ where $\phi(v_1, v_2, ...v_n)$ is a Δ_0^R formula. Each of Ax-1, Ax-2 and Ax-3 will contain a common set of nine Π_1^R axioms, called Q_0 and listed below. The main purpose of Q_0 will be to define the constructs of addition, multiplication, integer-successor, maximum and also $=$ and $\leq$.

$$1 \, = \, 0' \;\; \wedge \;\; 2 \, = \, 1' \;\; \wedge \;\; 0 \, = \, 0 \;\; \wedge \;\; 0' \neq 0 \;\; \wedge \;\; 0 \leq 0 \;\; \wedge \;\; \neg \, [\, 0' \, \leq \, 0 \,] \tag{5}$$

$$\forall \, x \;\; (\;\; x + 0 \, = \, x \quad \wedge \quad x \cdot 0 \, = \, 0 \quad \wedge \quad x \cdot 1 \, = \, x \;\;) \tag{6}$$

$$\forall x \; \forall y \;\; (\;\; x' \, = \, y' \quad \Longleftrightarrow \quad x \, = \, y \;\;) \tag{7}$$

$$\forall x \; \forall y \;\; (\;\; x \, \leq \, y \quad \Longleftrightarrow \quad (\, x' \, \leq \, y \; \vee \; x \, = \, y \;\;) \tag{8}$$

$$\forall x \; \forall y \qquad x \cdot y' \; = \; (x \cdot y) + x \quad \wedge \quad x + y' \; = \; (x + y)' \tag{9}$$

$$\forall x \; \forall y \; \forall z \quad [\, x = y \; \wedge \; y = z \,] \; \Rightarrow \; [\, x = z \; \wedge \; z = x \,] \tag{10}$$

$$\forall x \; \forall y \; \forall z \quad [\, x = y \; \wedge \; y \leq z \,] \; \Rightarrow \; x \leq z \tag{11}$$

$$\forall x \; \forall y \; \forall z \quad [\, x = y \; \wedge \; z \leq y \,] \; \Rightarrow \; z \leq x \tag{12}$$

$$\forall x \, \forall y \;\; (\; x \, \leq \, y \; \Rightarrow \; \mathrm{Max}(x, y) = y \;) \; \wedge \; (\; y \, \leq \, x \; \Rightarrow \; \mathrm{Max}(x, y) = x \;) \tag{13}$$

In the context of the above definition for Q_0, the Ax-1 and Ax-2 axiomatizations for $I\Sigma_0$ will be defined formally as the union of Q_0 with all instances of respectively Equations (1) and (2)'s induction schemas where $\phi(x, y)$ is a Δ_0 formula. Similarly, $I\Delta_0^R$ will be defined as the union of Q_0 with all instances of Equation (2)'s induction schemas where $\phi(x, y)$ is Δ_0^R.

This paragraph will define a set of Π_1^R sentences, called **Trivial-R**, that has the property that $I\Delta_0^R +$ Trivial-R proves the same set of theorems as the more conventional Ax-1 and Ax-2 axiomatization for $I\Sigma_0$. In our discussion, a tuple $(a_0, a_1, a_2, ...a_N)$ is called a **Split Representation** of an non-negative integer x when the following condition is satisfied:

$$x \;\; = \;\; \sum_{i=1}^{N} \, a_i \, * \, (a_0 + 1)^{i-1} \;\;\; \text{AND} \;\;\; a_1 \, \leq \, a_0 \, \wedge \, a_2 \, \leq \, a_0 \, \wedge \, ... \, a_N \, \leq \, a_0 \tag{14}$$

For a fixed integer N, let $\mathrm{Split}^N(\, x \, , \, a_0 \, , \, a_1 \, , \, ... \, a_N \,)$ denote a Δ_0^R formula indicating (14) is satisfied.

For each of the arithmetic operators of $+$, $*$, Max, $=$ and $\leq$, the axiom system Trivial-R will have available a family of Δ_0^R predicates and Π_1^R axioms for simulating the operations of these functions when they manipulate Split Representations of integers. Thus for a fixed triple (I, J, K), let $\mathrm{Mult}^{I,J,K}(a_0, a_1....a_I, b_0, b_1....b_J, c_0, c_1....c_K)$ designate a Δ_0^R predicate simulating the action of integer multiplication when its input is the two split integers of $(a_0, a_1....a_I)$ and $(b_0, b_1....b_J)$ and its resultant is the multiplicative product of $(c_0, c_1....c_K)$. The accompanying Π_1^R axiom of our system Trivial-R, that indicates this predicate operates correctly, will then be:

$$\forall\, x \;\; \forall\, y \;\; \forall\, z \;\; \forall\, a_0 \;\; \forall\, a_1 \,...\, \forall\, a_I \;\; \forall\, b_0 \;\; \forall\, b_1 \,...\, \forall\, b_J \;\; \forall\, c_0 \;\; \forall\, c_1 \,...\, \forall\, c_K$$

$$\{ \quad [\quad \text{Split}^I(x, a_0...a_I) \;\wedge\; \text{Split}^J(y, b_0...b_J) \;\wedge\; \text{Split}^K(z, c_0...c_K) \quad] \quad \Longrightarrow$$

$$[\;\; x * y = z \quad \Longleftrightarrow \quad \text{Mult}^{I,J,K}(a_0...a_I, b_0...b_J, c_0...c_K) \;\;] \quad \}$$

Likewise, Trivial-R will have available the suitable Π_1 analogs of the above axiom simulating similarly the formalisms of addition, maximum, equality, and less-than-or equals among split integers.

Henceforth, **Ax-3** will denote the axiom system $I\Delta_0^R$ + Trivial-R. Section 3 will prove that Ax-3 proves the same set of theorems as Ax-1 and Ax-2.

Definition 1: Let $\alpha \supseteq \beta$ denote that α's set of formal axioms includes all β's axioms. (This definition of "$\supseteq$", is stronger than the *more modest construct* that α proves all β's theorems.) Also assuming α denotes a **consistent** axiom system and D denotes a deductive method, (α, D) will be called a **Threshold** for the Second Incompleteness Effect iff all consistent extensions $\alpha^* \supseteq \alpha$ have the property that α^* is unable to prove the consistency of its proofs using deduction method D. Otherwise, (α, D) will be called an **Anti-Threshold**. (It means that *some consistent* $\alpha^* \supseteq \alpha$ can prove a theorem affirming its own consistency under deduction method D.)

In this context, our main result will be that Ax-3 is an anti-threshold for the Herbrandized version of the Second Incompleteness Theorem. This means that there must assuredly exist some *consistent* system $\alpha^* \supseteq$ Ax-3, where α^* can prove a theorem affirming its own Herbrand consistency.

This result is surprising because Ax-1 and Ax-2 are at the same time Herbrandized thresholds. We again remind the reader that logically equivalent systems can have opposite threshold properties. This is because $\alpha \cong \beta$ denotes merely that the two axiom systems α and β prove the same set of theorems. Under our notation, it does not imply that either α or β *can prove the* statement "$\alpha \cong \beta$". (This is the intuitive explanation for why Ax-3's threshold property will diverge from that of Ax-1 and Ax-2.)

3 Basic Framework and Underlying Intuition

This section will formally prove that Ax-3 proves the same set of theorems as Ax-1 and Ax-2. It will also *intuitively explain* why Ax-3's threshold property (under Definition 1) is different from that of Ax-1 and Ax-2.

Theorem 1 *Each of Ax-1, Ax-2 and Ax-3 prove the same set of theorems.*

Proof Sketch: It is well known Ax-1 and Ax-2 prove the same set of theorems. Thus to establish Theorem 1, we need only show Ax-2 ≅ Ax-3. Our proof will use the fact that Paris and Dimitracopoulos [20] have observed that in model-theoretic sense, there is a 1-to-1 correspondence between Δ_0 formulae and their equivalent representations in a Δ_0^R form.

In the interests of brevity, we will omit formally proving that Ax-2≅Ax-3. Instead, our proof-sketch will explore an example illustrating the underlying intuition behind this invariant.

Thus, let $\psi(x,y)$ denote a Δ_0^R formula. For any integer k, it is possible to construct a Δ_0^R formula $\psi^*(x, y_0, y_1, ...y_k)$ that is the counterpart of $\psi(x,y)$ for split representations of integers by satisfying the following property:

$$\forall\, x \;\; \forall\, y \;\; \forall\, y_0 \;\; \forall\, y_1 \; ... \; \forall\, y_k$$

$$\{ \;\; \text{Split}^k(y, y_0 ... y_k) \quad \Longrightarrow \quad [\; \psi(x,y) \quad \Longleftrightarrow \quad \psi^*(x, y_0, y_1, ...y_k) \;\;] \;\; \} \tag{15}$$

Let $\text{Size}_L(y_0, y_1, ...j_k)$ denote a Δ_0^R formula indicating that $(y_0, y_1, ...j_k)$ represents an integer $\leq\; L$. Then Ax-3 can use its Trivial-R axioms to first prove Equation (15), and then to formally prove that the two Δ_0 formulae of $\exists y \leq x^k \;\; \psi(x,y)$ and $\forall y \leq x^k \;\; \psi(x,y)$ are equivalent to the respective Δ_0^R formulae of:

$$\exists\, y_0 \;\leq\; x \;\, \exists\, y_1 \;\leq\; x \; ... \; \exists\, y_k \;\leq\; x \qquad \text{Size}_{x^k}(y_0, y_1, ...y_k) \wedge \psi^*(x, y_0, y_1, ...y_k)$$

$$\forall\, y_0 \;\leq\; x \;\, \forall\, y_1 \;\leq\; x \; ... \; \forall\, y_k \;\leq\; x \qquad \text{Size}_{x^k}(y_0, y_1, ...y_k) \Rightarrow \psi^*(x, y_0, y_1, ...y_k)$$

Thus by essentially applying n iterations of this technique (and its obvious analogs) to any initial Δ_0 formula with n bounded quantifiers, Ax-3 can transform an arbitrary Δ_0 formula into a provably equivalent Δ_0^R formula. It thus follows that although the Ax-3 system contains technically only instances of Equation (2)'s axiom schema for Δ_0^R formulae, it nevertheless has an ability to formally prove as theorems all the remaining instances of this axiom schema for Δ_0 formulae *as well.* □

Our proof that Ax-3 is an anti-threshold for the Herbrandized version of the Incompleteness Theorem will appear in Section 4. Before starting that discussion, the underlying intuition as to why Ax-2 and Ax-3 do operate so very differently should be explained.

Let Υ_n denote the Δ_0 sentence defined by Equation (16). Note this sentence is comprised of $O(\; n\;)$ logic symbols, and it asserts that the variables $v_0\,,\; v_1\,,\; v_2\,,\; ...v_n\,,$ have the properties that $v_i \;=\; 2^{2^i}$.

$$\exists\; v_0 \;\leq\; 2 \;\;\; \exists\; v_1 \;\leq\; v_0 * v_0 \;\;\; \exists\; v_2 \;\leq\; v_1 * v_1 \;\;\; ... \;\;\; \exists\; v_n \;\leq\; v_{n-1} * v_{n-1}$$

$$v_0 \;=\; 2 \;\wedge\; v_1 \;=\; v_0 * v_0 \;\wedge\; v_2 \;=\; v_1 * v_1 \;\wedge\; ... \wedge v_n \;=\; v_{n-1} * v_{n-1} \tag{16}$$

It is easy to see there exists some Δ_0^R sentence, called say Υ_n^R that is the counterpart of Equation (16) written in a notation using split integers. This sentence will indicate the existence of a sequence of split integers $S_0, S_1, S_2, ...S_n$, where S_i represents the quantity 2^{2^i}.

However although they in some sense represent equivalent concepts, there is a fundamental difference between the Δ_0 sentence Υ_n and its Δ_0^R counterpart Υ_n^R. This difference is easiest to explain if one uses a logical language that has only 3 named constants, 0, 1 and 2, and if split integers are encoded as base 2 numbers. Then Υ_n^R will be encoded as a sequence of at least 2^n characters, but Υ_n's length has a sharply different O(n) magnitude. As a consequence of this distinction (and its generalizations), we can establish that although Ax-2 and Ax-3 prove the identical set of formal theorems, their proofs of many theorems can differ by an exponential magnitude in length

This fact is crucial for understanding why these two formalisms have different incompleteness-threshold properties. It explains intuitively why Ax-2 (in our companion paper [40]) obeys the Second Incompleteness Theorem, but Ax-3 is shown (in the next section) to actually evade it.

4 Main Analysis

A sentence ψ in the propositional calculus will be called an **Anti-Tautology** iff ψ is unsatisfiable (i.e. $\neg\,\psi$ is a tautology). Our definition of Herbrand deduction will be identical to the definitions used by Adamowicz, Hájek-Pudlák and Kołodziejczyk, [1, 10, 15], except that we will use a dual version of this definition that follows from De Morgan's Rule, where disjunctions are replaced with conjunctions and where tautologies are replaced with anti-tautologies. In other words, our definition will use the well-known identity that

$$\bigvee_{i=1}^{n} \neg\,\phi_i \quad = \quad \neg \bigwedge_{i=1}^{n} \phi_i \tag{17}$$

Our definition of Herbrand deduction will differ from its more conventional definitions by using the right (instead of left) side of (17)'s identity. This change in notation is unnecessary, but it does help simplify our proofs.

Let Ψ denote an arbitrary prenex normal sentence such as the prototype below, whose open subcomponent is denoted as $\widetilde{\psi}$.

$$\forall\, x_1 \;\exists\, y_1 \;\forall\, x_2 \;\exists\, y_2 \;.... \quad \forall\, x_n \;\exists\, y_n \quad \widetilde{\psi}\,(x_1, y_1 ... x_n, y_n) \tag{18}$$

In a context where $f_1^\psi(x_1),\; f_2^\psi(x_1,x_2) \;...\; f_n^\psi(x_1,x_2,...x_n)$ are new function symbols, Equation (19) is called the Skolemization of Equation (18).

$$\forall x_1 \,\forall x_2 \;....\; \forall x_n \;\; \widetilde{\psi}\;[\; x_1 \,,\, f_1^\psi(x_1)\,,\, x_2 \,,\, f_2^\psi(x_1,x_2) \,...\, x_n \,,\, f_n^\psi(x_1,x_2...x_n)\,] \tag{19}$$

In a context where L is a logical language and α is an axiom system, we will let C_L and F_L denote the set of constant and function symbols associated with L. Similarly, F_α will denote the set of "Skolemized" function symbols associated with α's axioms. Thus using (18) and (19)'s notation, let α denote a system of axioms Ψ_1, Ψ_2, Ψ_3..., and for an arbitrary index i let its Skolemized function symbols carry names such as $f_i^{\psi_1}$, $f_i^{\psi_2}$, $f_i^{\psi_3}$, ... The **Herbrandized Terms** for this ordered pair (α, L) will then be defined to be the set of all terms generated by the constants from the set C_L combined with the functional operations from the set $F_\alpha \cup F_L$.

A **Herbrandized Instance** of a Skolemized axiom is a sentence identical to this axiom except that all its universally quantified variables are replaced by Herbrandized terms. For instance in a context where $T_1, T_2, T_3...$ are Herbrandized terms, Equation (20) is such an instance of (19)'s axiom:

$$\widetilde{\psi}\ [\ T_1\ ,\ f_1^\psi(T_1)\ ,\ T_2\ ,\ f_2^\psi(T_1,T_2)\ ...\ T_n\ ,\ f_n^\psi(T_1,T_2...T_n)\ ] \tag{20}$$

Let $\perp$ denote the logical constant of FALSE. A **Herbrandized Proof** of $\perp$ from the axiom system α is defined as a finite collection of Herbrandized instances of α together with a proof, in the pure propositional calculus, that the conjunction of these instances is an anti-tautology.

Definition 2 : Using our revised notation convention, the theorem Υ will be said to have a **Herbrandized Proof** from the axiom system β if and only if the union of the axiom system β with the added sentence $\neg\Upsilon$ produces a Herbrandized proof of $\perp$.

More Notation: Let us say that a function $G(x_1, x_2, ... x_n)$ is a **Non-Growth Function** iff $G(x_1, x_2, ... x_n) \leq \text{Max}(x_1, x_2, ... x_n)$. Define a set S of functions to be an **Arithmetic Controlled Set** iff S includes the arithmetic functions of addition, multiplication and successor and all its other functions are non-growth functions. Also, define a term t to be an **Arithmetically Controlled Term** iff t is a term that uses only the symbols of 0, 1 and 2 as its inputted constants and all its function symbols come from some Arithmetic Controlled Set S. Thus if G_1 and G_2 are non-growth functions, Equation (21) represents an arithmetically controlled term.

$$G_1[\ (1+1)*(1+1)\ ,\ 1+0\ ]\ *\ G_2(\ 1+1+0\ ,\ 1+1+1+1\ ) \tag{21}$$

Also, in a context where C_t and F_t denote the number of constant and function symbols in t, we will use the following notation:

1. MinG(t) will denote the quantity $2^{C_t+F_t}$.

2. Val(t) will denote the quantity represented by the term t.

For example if $G_1(x,y) = |x-y|$ and $G_2(x,y) = \text{Min}(x,y)$ then Equation (21)'s term t will have $Val(t) = 3*4 = 12$ and MinG(t) $= 2^{25}$ (because t contains 12 function symbols and 13 constant symbols).

Lemma 1 *Let t be an arithmetically controlled term which satisfies the inequality $Val(t) \geq 4$. Then $Val(t) < MinG(t)$*

Proof Sketch: Suppose for some $k \geq 2$, that Val(t) $= 2^k$. Then it easy to see that t's maximally compressed representation as an *arithmetically controlled term* is " $2*2*....2$ ". Thus MinG(t) $= 2^{2k-1} >$ Val(t) $= 2^k$ is valid in this case because the preceding product has k appearances of the constant 2 connected by $k-1$ appearances of the multiplication symbol. Moreover, it is easily proven that terms, which are not powers of 2, are never represented in a more compressed form than the greatest power of 2 that they exceed. Thus Lemma 1 is valid for all terms where Val(t) ≥ 4. □

Definition 3. For a fixed constant $B > 0$, a set S of functions is defined to be a $B-$**Bounded Arithmetic Set** iff S includes the arithmetic functions of addition, multiplication and successor and all its other functions G satisfy the constraint that

$$G(x_1, x_2, ...x_n) \leq \text{Max}(x_1, x_2, ...x_n) \text{ when } \text{Max}(x_1, x_2, ...x_n) < B \tag{22}$$

Also, we will say a term t is a **B-Bounded Arithmetic Term** iff t is a term that uses only the symbols of 0, 1 and 2 as its inputted constants and all its function symbols come from some B-Bounded Arithmetic Set S. Lemma 2. provides the generalization of Lemma 1 for B-bounded arithmetic terms. Its proof is omitted because it is similar to Lemma 1's proof.

Lemma 2 *Suppose that t is a $B-$bounded arithmetic term with $MinG(t) < B$ and $Val(t) \geq 4$. Then $Val(t) < MinG(t)$*

Definition 4. Let Φ denote the Π_1^R sentence below whose Δ_0^R subformula is defined by $\widetilde{\phi}(\ a_1, a_2\ ...\ a_n\ )$.

$$\forall a_1 \forall a_2 ... \forall a_n\ \widetilde{\phi}(a_1, a_2 ... a_n) \tag{23}$$

For any $B \geq 1$, Equation (23) is called a $B-$**Bounded Valid** Π_1^R **sentence** iff (24) is valid under the standard model of the natural numbers:

$$\forall\ a_1 < B\ \forall\ a_2 < B\ ...\ \forall\ a_n < B \quad \widetilde{\phi}(\ a_1, a_2\ ...\ a_n\ ) \tag{24}$$

Definition 5. An axiom system α will be said to satisfy the **Canonical Arithmetic Condition** when all α's axioms are Π_1^R sentences and they include Q_0's nine axioms (i.e. Equations (5)–(13)).

Definition 6. Let Θ denote a methodology for assigning Gödel numbers to Herbrand proofs (which are henceforth denoted as P). Let us recall that MinG(t) was defined by Item (i) in this section. Define Θ to be a **Conventional Encoding Method** if $\Theta(P) > \text{MinG}(t)$ whenever the proof P contains the Herbrand term t . (Such encodings are called "conventional" because all usual methods for encoding Herbrand proofs satisfy $\Theta(P) > \text{MinG}(t)$.)

Theorem 2 *Suppose α is a canonical arithmetic axiom system consisting of $B-$Bounded Valid Π_1^R sentences and Θ again satisfies Definition 6's Conventional Encoding property. Then any Herbrand proof P of $\perp$ from the axiom system α will satisfy the inequality that $\Theta(P) > B$.*

General Comments about Theorem 2 and its Proof: At an intuitive level, Theorem 2 can be viewed as a consequence of the machineries of Lemma 2 and Definitions 4-6. This is because the B-Bounded validity condition in Theorem 2's hypothesis can be used to show that a Herbrand proof P of $\perp$ must contain some term t where $\text{Val}(t) \geq B$. In this context, the combination of Lemma 2 and Definition 6 will imply that such a term will force P 's Gödel number to exceed the lower bound of B.

A formal proof of Theorem 2 is available in the Appendix. Our recommendation is that if a reader does wish to examine this appendix's proof, he do so only after he finishes the next two pages of this article. They will explain how the formalism of Theorem 2 shall enable us to prove the surprising result that the Ax-3 axiomatization for $I\Sigma_0$ is an anti-threshold for the Herbrandized version of the Second Incompleteness Theorem.

Theorem 3 *For any arbitrary axiom system α and deduction method D, let Diagonal(α, D) denote the following sentence:*

> Diagonal(α, D) = There is no proof (using deduction method D) of the "falsity sentence" $\perp$ from the union of the axiom system α with *this* sentence "Diagonal(α, D) " (looking at itself).

Also, in a context where i = 1, 2 or 3, let Diag(i) denote the special variant of Diagonal(α, D) where α = Ax$-i$ and D designates Herbrand deduction. Both these constructs are well defined, and Diag(i) also has a Π_1^R encoding.

Sketch of Theorem 3's proof and comment about its significance. As early as 1938, Kleene observed [14] that a form of the sentence Diagonal(α, D) was well defined. More recently, Willard [34, 37] observed this sentence also has a Π_1 encoding in the conventional language of arithmetic. It is straightforward to generalize [34, 37]'s result to establish that Diag(i) has a well defined Π_1^R encoding (thus completing Theorem 3's proof.) □

Clarifying Comment: One should be somewhat cautious in interpreting the meaning of Theorem 3. It does not indicate that Diag(i) is a logically valid statement under the standard model of the natural numbers. Rather, it merely indicates Diag(i) is a well defined Π_1^R sentence. In fact, Diag(1) and Diag(2) can be proven to be logically invalid statements (see footnote [1]). In contrast, Theorem 4 (below) will prove Diag(3) is logically valid.

Theorem 4 *Let Ax-3* denote the union of Ax-3 with the sentence Diag(3). Then Ax-3* is consistent. (Hence it follows that Ax-3 is an "anti-threshold" for the Herbrandized version of the Second Incompleteness Theorem under Definition 1's notation convention.)*

Proof of the Consistency Property of Ax-3* : Suppose for the sake of establishing a proof-by-contradiction that Ax-3* was inconsistent. Then one could identify a proof P of $\perp$ whose Gödel number $\Theta(P)$ is the smallest Gödel number of a Herbrand proof of $\perp$ from Ax-3*. We will now construct from P an alternate Herbrand proof R of $\perp$ where $\Theta(R) < \Theta(P)$. The formal construction of such a R will suffice for our proof by contradiction to reach its desired end because such a R cannot possibly exist (on account of P's minimality property).

Our strategy is to use Theorem 2 to construct R from P . Theorem 2 is relevant to Ax-3* (but not also to Ax-1's or Ax-2's analogs of it) because only all the formal axioms of Ax-3* are assuredly Π_1^R sentences. This distinction arises because the induction schemes for Ax-3 (and thus Ax-3*) uses Δ_0^R formulae (unlike the more liberal Ax-1 and Ax-2 induction schemes that replace Δ_0^R formulae with the less manageable Δ_0 expressions)

On account of the fact that all Ax-3* 's axioms are Π_1^R sentences, we may apply Theorem 2 to conclude that for some $B < \Theta(P)$, at least one of the axiom sentences of Ax-3* fail to be a B-Bounded valid Π_1^R sentence. Moreover, it is obvious that all the axioms of Ax-3 possess an unbounded level of validity (i.e they are $B-$Bounded valid *for all possible B.*) Hence, these two observations imply Diag(3) fails to be $B-$bounded valid (simply because some axiom from Ax-3* must fail to be $B-$bounded valid, and Diag(3) is the only axiom belonging to Ax-3* that is not also a member of Ax-3.)

The latter observation, combined with Diag(3)'s definition, implies that some R with $\Theta(R) < B$ must be another proof of $\perp$. (This is because Diag(3)'s failure to be $B-$bounded

[1]For an arbitrary axiom system α, let α^D denote the union of α with the added sentence Diagonal(α, D). Most such systems α^D are known to be inconsistent because they would otherwise violate Gödel's Second Incompleteness Theorem. The main point of our prior research [32, 34, 37, 39] is that the usual paradigm where an essentially classic Gödel-like diagonalization argument will render α^D inconsistent applies to most, *but not all* systems α^D . Thus, it turns out that the classic Gödel-like paradigm applies to Ax-1 and Ax-2 under Herbrand deduction. On the other hand, the final result of this paper (Theorem 4) will prove that Ax-3 is quite different.

valid implies such an R must assuredly exist.) Hence $\Theta(R) < B < \Theta(P)$ and our proof-by-contradiction is finished because P's previously presumed minimality has been contradicted by R. $\square$

5 Concluding Remarks

The research reported here is essentially the third facet of a 3-part project. The first facet was our year-2002 JSL article [35]. It established that the main textbook axiomatization of $I\Sigma_0$ [10, 13, 17], which we have called Ax-1, satisfies the semantic tableaux and Herbrandized versions of the Second Incompleteness Theorem. The second part of this project [40] had generalized the preceding incompleteness result so that it also applied to Ax-2. This current article has shown that Ax-3, quite surprisingly, evades the Herbrandized version of the Second Incompleteness Theorem.

One reason our results are of interest is because there have been no prior examples in the literature where a natural axiom system, such as $I\Sigma_0$, can have several equivalent axiomatizations, some of which satisfy the Herbrandized form of the Second Incompleteness Theorem and others of which represent what Definition 1 calls its "anti-thresholds". (As we noted earlier, it is possible for two logically equivalent axiom systems, α and β, to have opposite threshold-incompleteness properties because the fact that $\alpha \cong \beta$ does not imply that either of these two systems can formally prove " $\alpha \cong \beta$".)

The many generalizations of the Second Incompleteness Theorem are clearly significantly more important than its occasional boundary-case exceptions. Nevertheless, these partial exceptions to the Second Incompleteness Theorem should not be ignored. Gödel's Incompleteness Theorem is usually regarded as the paramount discovery of 20th century mathematics. It thus beckons the academic community to explore its possible boundary case exceptions, so that an understanding of its full meaning can be sharpened and made more precise. Within such a limited-but-precise framework, the anomalous behavior of Ax-3, documented in this article, should be of scholarly interest.

Acknowledgment: I warmly thank Leszek Kołodziejczyk for the interesting questions that he had emailed me on November 16, 2005. Those questions played a major role in encouraging me to re-examine my earlier work in [35] for one more time. I also thank Zofia Adamowicz and Konrad Zdanowksi, who helped at least partially stimulate the emailed questions from L. Kołodziejczyk through their conversations with him in Warsaw. Those questions from Leszek Kołodziejczyk had greatly stimulated my research in this article.

6 Appendix: The Proof For Theorem 2

Our proof of Theorem 2 will essentially be an easy consequence of the machineries of Definition 7 and of two further lemmas.

Definition 7. Consider the possibility that Ψ is the prenex normal sentence, whose open part is formalized by $\widetilde{\psi}\,(,x_1,y_1...x_n,y_n)$, shown in Equation (25) and whose Skolemized normalized form is illustrated by Equation (26).

$$\forall\ x_1\ \ \exists\ y_1\ \ \forall\ x_2\ \ \exists\ y_2\ \ ....\quad \forall\ x_n\ \ \exists\ y_n\quad \widetilde{\psi}\,(,x_1,y_1...x_n,y_n) \tag{25}$$

$$\forall x_1\ \forall x_2\ ....\ \ \forall x_n\quad \widetilde{\psi}\,[\,x_1\,,\,f_1^{\psi}(x_1)\,,\,x_2\,,\,f_2^{\psi}(x_1,x_2)\,...\,x_n\,,\,f_n^{\psi}(x_1,x_2...x_n)\,] \tag{26}$$

For any $B\ \geq\ 1$, Equations (25) and (26) will be called a $B-$**Bounded Good Skolemization** iff one can define (26)'s Skolem functions f_1^{Ψ} , f_2^{Ψ} ... f_n^{Ψ} so that they simultaneously satisfy Definition 3's $B-$Bounded requirement and Equation (27) under the standard model of the natural numbers.

$$\forall\ x_1\ <\ B\ \ \forall\ x_2\ <\ B\ \ ....\quad \forall\ x_n\ <\ B$$

$$\widetilde{\psi}\ [\ x_1\,,\,f_1^{\psi}(x_1)\,,\,x_2\,,\,f_2^{\psi}(x_1,x_2)\ ...\ x_n\,,\,f_n^{\psi}(x_1,x_2...x_n)\ ] \tag{27}$$

Likewise, we will say an axiom system α has a $B-$**Bounded Good Skolemization** iff all its axioms are so Skolemized.

Lemma 3 *Using the notation conventions from Definitions 4 and 7, every $B-$Bounded Valid Π_1^R sentence can be rewritten into a logically equivalent form that has a $B-$Bounded Good Skolemization.*

Proof. Follows immediately from the definitions of Bounded Validity and Bounded Good Skolemizations (i.e. see Definitions 4 and 7).

Lemma 4 *Using the notation conventions from Definitions 5-7, suppose that α is a canonical arithmetic system consisting of prenex sentences which possess $B-$Bounded Good Skolemizations and that Θ satisfies the Conventional Encoding property. Then any Herbrand proof P of $\perp$ from the axiom system α will satisfy $\Theta(P)\ >\ B$.*

Proof-by-contradiction: Consider the contrary possibility that the inequality $\Theta(P)\ >\ B$ failed and that P is a Herbrand-proof of $\perp$ from the axiom system α where $\Theta(P)\ \leq\ B$.

* **

Definition 6 had indicated every term T in the proof P satisfies $\Theta(P) > \text{MinG}(T.)$ Also, Lemma 2 implied Val(T) <MinG(T). These inequalities and *** imply that every term T in the proof P satisfies

$$\text{Val}(T) \ < \ B \tag{28}$$

Equation (28) implies all the terms $T_1, T_2, T_3...$ in the Herbrandized instances in the proof P satisfy Val$(T_i) < B$. The normalized form of an instance of a Skolemized axiom is illustrated by Equation (29). The combination of our Val$(T_i) < B$ inequalities together with (27)'s $B-$Bounded constraint on α's axioms implies that *each such instance of (29) appearing in the proof* P *must* **be automatically valid** *under the standard model of the natural numbers.*

$$\widetilde{\psi} \ [\ T_1 \ , \ f_1^{\psi}(T_1) \ , \ T_2 \ , \ f_2^{\psi}(T_1, T_2) \ ... \ T_n \ , \ f_n^{\psi}(T_1, T_2...T_n) \ ] \tag{29}$$

The latter observation completes our proof-by-contradiction because it essentially contradicts the initial statement $***$ that had started our contradiction proof. More precisely $***$ had asserted that P was a Herbrand-proof of $\perp$ from the axiom system α. However, the Footnote [2] shows that such is impossible when the last sentence of the preceding paragraph had indicated that each instance of (29)'s Skolemized axiom is actually fully valid under the standard model of the natural numbers. $\square$

Finishing the Proof for Theorem 2. It is easy to combine the machineries of Lemmas 3 and 4 to complete the proof of Theorem 2. This is because Lemma 3 had indicated that every $B-$Bounded Valid Π_1^R sentence can be rewritten into a logically equivalent form that has a $B-$Bounded Good Skolemization. Thus, Theorem 2 follows by simply taking such rewritten forms of α's axioms and then applying Lemma 4's machinery. $\square$

References

[1] Z. Adamowicz, "Herbrand Consistency and Bounded Arithmetic", *Fundamenta Mathematica* 171 (2002) 279-292.

[2] Z. Adamowicz and T. Bigorajska, "Existentially Closed Structures and Gödel's Second Incompleteness Theorem", *Journal of Symbolic Logic* 66 (2001), 349-356.

[3] Z. Adamowicz and P. Zbierski, "On Herbrand consistency in weak theories", *Archive for Mathematical Logic* 40 (2001) pp. 399-413.

[4] A. Bezboruah and J. Shepherdson, "Gödel's Second Incompleteness Theorem for Q", *Journal of Symb Logic* 41 (1976) 503-512.

[2]The point here is simply that a conjunction of Skolemized instances can produce a proof of $\perp$ only when there exists no model M where all these instances are simultaneously valid. Hence when the preceding paragraph shows that all these Skolemized instances are simultaneously valid under the Standard Model of the Natural Numbers, it implies that certainly *no such proof of* $\perp$ *can feasibly exist !*

[5] S. Buss and A. Ignjatovic, “Unprovability of Consistency Statements in Frag. of Bounded Arithmetic”, *Annals Pure and Applied Logic* 74 (1995)221-244.

[6] M. Davis, *Computability and Unsolvability* McGraw Hill 1983.

[7] M. Davis, H. Putnam and J. Robinson, “The Decision Problem for Exponential Diophantine Equations”, *Annals of Mathematic* 74 (1961) pp. 425-436.

[8] S. Feferman, “Arithmetization of Mathematics in a General Setting”, *Fundamenta Mathematica* 19 (1960) pp. 35-92.

[9] K. Gödel, “ Über formal unentscheidbare Sätse der Principia Mathematica und Verwandte Systeme I”, *Monatshefte für Math. Phys.* 37 (1931) pp. 349-360.

[10] P. Hájek and P. Pudlák, *Metamathematics of First Order Arithmetic,* Springer Verlag 1991.

[11] J. Herbrand, Recherches sur la theorie de la demonstration, *Travaux de la Soceite et des Lettres de Varsvoie* III 33 pp. 33-160 (1930).

[12] D. Hilbert and P. Bernays, *Grundlagen der Mathematic*, Springer 1939.

[13] R. Kaye, *Models of Peano Arithmetic*, Oxford University Press, 1991.

[14] S. Kleene, “On the Notation of Ordinal Numbers”, *Journal of Symbolic Logic* 3 (1938), pp. 150-156.

[15] L.A. Kołodziejczyk, On the Herbrand Notion of Consistency for Finitely Axiomatizable Fragments of Bound Arith, to appear in *Journal of Symbolic Logic*

[16] J. Krajícek, “A Note on Proofs of Falsehood”, *Archive for Mathematical Logic* 26 (1987) 169-176.

[17] J. Krajícek, *Bounded Propositional Logic and Complexity Theory,* Cambridge University Press, 1995.

[18] M. H. Löb, A Solution to a Problem by Leon Henkin, *Journal of Symbolic Logic* 20 (1955) pp. 115-118.

[19] Y. Matiyasevich, *Hilbert's Tenth Problem*, MIT Press (1993)

[20] J. Paris and C. Dimitracopoulos, “Truth definitions for Δ_0 formulae”, *Monagraphie de L'Enseignement Mathematique* 30 (1982) pp. 317-329.

[21] J. Paris and A. Wilkie, “Δ_0 Sets and Induction”, *1981 Jadswin Conference Proceedings,* (Leeds University Press) pp. 237-248.

[22] P. Pudlák, “On Lengths of Proofs of Finisitic Consistency Statements in First order Theories”, *Logic Colloquium 84*, North Holland (1986) pp. 165-196.

[23] P. Pudlák, “Cuts, Consistency Statements and Interpretations”, *Journal of Symbolic Logic* 50 (1985) 423-442

[24] P. Pudlák, “On the Lengths of Proofs of Consistency”, in *Collegium Logicum: 1996 Annals of the Kurt Gödel Society* (Vol 2), Springer-Wien-NewYork, pp. 65-86.

[25] C. Smoryński, "Non-standard Models and Related Developments in the Work of Harvey Friedman", in *Harvey Friedman's Research in Foundations of Math.*, North Holland 1985, pp. 179-220.

[26] R. Solovay, Telephone conversations between Solovay and Willard (during April of 1994) describing Solovay's generalization of one of Pudlák's theorems [23], using some of Nelson's and Wilkie-Paris's methodologies. The Appendix A of [34]offers a 4-page interpretation of the fundamental intuition behind Solovay's idea.

[27] G. Takeuti, "Gödel Sentences of Bounded Arithmetic", *Journal of Symbolic Logic* 65 (2000) pp. 1338-1346

[28] A. Tarski, A. Mostowski and R. M. Robinson, *Undecidable Theories*, North Holland, 1953.

[29] A. Visser, "The Unprovability of Small Inconsistency", *Archive for Mathematical Logic* 32 (1993) pp. 275-298.

[30] A. Visser, "Faith and Falsity", *Annals Pure & Applied Logic* 131 (2005), pp. 103-131.

[31] A. Wilkie and J. Paris, "On the Scheme of Induction for Bounded Arithmetic", *Annals of Pure and Applied Logic* (35) 1987, 261-302

[32] D. Willard, "Self-Verifying Axiom Systems", *Third Kurt Gödel Colloq* (1993), Springer-Verlag LNCS#713, 325-336.

[33] D. Willard, "The Semantic Tableaux Version of the Second Incompleteness Theorem Extends Almost to Robinson's Arithmetic Q", Springer Verlag LNCS#1847, July 2000, pp. 415-430.

[34] D. Willard, "Self-Verifying Systems, the Incompleteness Theorem and the Tangibility Principle", in *Journal of Symbolic Logic* 66 (2001) pp. 536-596.

[35] D. Willard, "How to Extend The Semantic Tableaux And Cut-Free Versions of the Second Incompleteness Theorem Almost to Robinson's Arithmetic Q", in *Journal of Symbolic Logic* 67 (2002) pp. 465–496.

[36] D. Willard, "A Version of the Second Incompleteness Theorem For Axiom Systems that Recognize Addition But Not Multiplication as a Total Function", *First Order Logic Revisited,* Logos Verlag (Berlin) 2004, pp. 337–368.

[37] D. Willard, "An Exploration of the Partial Respects in which an Axiom System Recognizing Solely Addition as a Total Function Can Verify Its Own Consistency", *Journal of Symbolic Logic* 70 #4 (2005) pp. 1171-1209.

[38] D. Willard, "On the Partial Respects in which a Real Valued Arithmetic System Can Verify its Tableaux Consistency" *Automated Reasoning with Analytic Tableaux (2005).* SpringerVerlag LNCS#3702, pp. 292-306.

[39] D. Willard, "A Generalization of the Second Incompleteness Theorem and Some Exceptions to It". To appear in 2006 in *Annals of Pure and Applied Logic.*

[40] D. Willard, "Passive Induction and a Solution to a Paris-Wilkie Open Question", an article submitted for publication and also available as Univeristy of Albany Computer Science Tehcnical Report SUNYA-CS-06-01 (February 1, 2006). This paper essentially generalizes [35]'s $I\Sigma_0$ version of the Second Incompleteness from Ax-1's definition of $I\Sigma_0$ to Ax-2's definition of it.