On the Broader Epistemological Significance of Self-Justifying Axiom Systems

Dan E. Willard

University at Albany Computer Science and Mathematics Departments

Abstract. This article will be a continuation of our research into self-justifying systems. It will introduce several new theorems (one of which will transform our previous infinite-sized self-verifying logics into formalisms or purely finite size). It will explain how self-justification is useful, even when the Incompleteness Theorem clearly limits its scope.

1 Introduction

Gödel's Incompleteness Theorem has two parts. Its first half indicates no decision procedure can identify arithmetic's true statements. Its "Second Incompleteness" result specifies sufficiently strong logics *cannot* verify their own consistency. Gödel was careful to insert a caveat into his historic paper [11], indicating a *diluted* form of Hilbert's Consistency Program might have some success:

$*$ *"It must be expressly noted Proposition XI represents no contradiction of the formalistic standpoint of Hilbert. For this standpoint presupposes only the existence of a consistency proof by finite means, and there might conceivably be finite proofs which cannot be stated in P or in ..."*

Some scholars have interpreted $*$ as, possibly, anticipating attempts to confirm Peano Arithmetic's consistency, via either Gentzen's formalism or Gödel's Dialetica interpretation. On the other hand, the Stanford's Encyclopedia's entry about Gödel quotes him, in its Section 2.2.4, stating he was hesitant to view the Second Incompleteness Theorem as fully ubiquitous, until learning of Turing's work. Moreover, Yourgrau [45] states von Neumann *"argued against Gödel himself"* in the early 1930's, about the definitive termination of Hilbert's consistency program, which *"for several years"* after [11]'s publication, Gödel *"was cautious not to prejudge"*. Also, it is known [6,13,45] that Gödel did initially presume the second theorem was false, before proving its stunning result.

In any case several year after he wrote $*$'s initial statement, Gödel gave a 1933 lecture [12], where he told his audience that Hilbert's initial 1926 objectives, summarized formally by $**$ below, had *"unfortunately"* no *"hope of succeeding along"* its originally intended plans.

$**$ (Hilbert [17] 1926): *"Where else would reliability and truth be found if even mathematical thinking fails? The definitive nature of the infinite has become necessary, not merely for the special interests of individual sciences, but rather for the honor of human understanding itself."*

U. Kohlenbach et al. (Eds.): WoLLIC 2014, LNCS 8652, pp. 221–236, 2014.
© Springer-Verlag Berlin Heidelberg 2014

Our research, in both the current article and prior papers [35–44], was stimulated by the prospect that we find $**$ enticing, even though the Second Incompleteness Theorem *unequivocally* demonstrates that logics *cannot* recognize their own consistency *in a robust sense.* Accordingly, we have studied *both* generalizations and boundary-case exceptions for the Second Incompleteness Theorem in [35–44]. The current article will seek to *both* strengthen these prior results, in the context of axiom systems with *strictly finite cardinalities*, and to also provide a more intuitive explanation of the meaning behind [35–44]'s results.

The thesis of this article will be delicate because there can be no doubt that the Second Incompleteness Theorem is sharply robust, when viewed from a conventional purist mathematical perspective. On the other hand, we will argue that there are certain facets of a "Self-Justifying Logics", that are tempting under a hard-nosed engineering perspective, contemplating sharply *curtailed forms* of Hilbert's goals. These results will be fragile *but not fully immaterial.*

2 Background Setting

Let (α, d) denote any axiom system and deduction method satisfying the simple **"Split Rule"** below [1]. This pair will be called **"Self Justifying"** when:

i one of α's theorems will state that the deduction method d, applied to the system α, will produce a consistent set of theorems, and
ii the axiom system α is in fact consistent.

For any (α, d), it is easy to construct a second $\alpha^d \supseteq \alpha$ that satisfies the Part-i requirement. For instance, α^d could consist of all of α's axioms plus an added **"SelfRef**(α, d)**"** sentence, defined as stating:

- There is no proof (using d's deduction method) of $0 = 1$ from the *union* of the system α with *this* sentence "SelfRef(α, d)" (looking at itself).

Kleene [20] noted how to encode rough analogs of "SelfRef(α, d)". Each of Kleene, Rogers and Jeroslow [19,20,29] noted α^d may, however, be inconsistent (despite SelfRef(α, d)'s assertion), thus causing it to violate Part-ii's requirement.

This problem arises in many contexts besides Gödel's paradigm, where α was an extension of Peano Arithmetic (see [1–5,7,9,11,14–16,18,21–23,25–34,38,39, 43]). Such results formalize paradigms where self-justification is infeasible, due to diagonalization issues. (It should, perhaps, be added that among this lengthy list of articles, it was especially [1, 4, 11, 23, 27, 31, 34]'s incompleteness results that influenced our work in [35–44].) In any case, the main point is that most logicians have hesitated to employ an analog of a SelfRef(α, d) axiom because $\alpha^d = \alpha$+SelfRef(α, d) is typically inconsistent.

[1] Our "Split Rule" is the trivial requirement that all the axiom sentences in α are technically *proper axioms*, and that deduction method d is required to include **BOTH** a finite number of rules of inference and whatever "logical axioms" are needed *(if any ?)* by d's methodology. (This trivial notation convention is helpful.)

Our research in [35,37,40–42] focused on paradigms where self-justification is feasible. It involved weakening the properties a logic can prove about addition and/or multiplication (to avoid potential difficulties). To be more precise, let $Add(x, y, z)$ and $Mult(x, y, z)$ denote 3-way predicates specifying $x + y = z$ and $x * y = z$. Then a logic will be said to **recognize** successor, addition and multiplication as **Total Functions** iff it includes sentences 1-3 as axioms.

$$\forall x \; \exists z \quad Add(x, 1, z) \tag{1}$$
$$\forall x \; \forall y \; \exists z \quad Add(x, y, z) \tag{2}$$
$$\forall x \; \forall y \; \exists z \quad Mult(x, y, z) \tag{3}$$

A logic α will be called **Type-M** iff it contains 1-3 as axioms, **Type-A** iff it contains only (1) and (2) as axioms, **Type-S** iff it contains only (1) as an axiom, and **Type-NS** iff it contains none of these axioms. The relationship of these constructs to self-justification is explained by items (a) and (b):

a. The existence of Type-A systems that can recognize their own consistency under semantic tableaux deduction, while proving analogs of all Peano Arithmetic's Π_1 theorems (in a slightly different language), was demonstrated in [40]. Also, [37,41] noted that some specialized forms of Type-NS systems can likewise recognize their own Hilbert consistency.

b. The above evasions of the Second Incompleteness Theorem are known to be near-maximal in a mathematical sense. This is because the combined work of Pudlák, Solovay, Nelson and Wilkie-Paris [24,27,31,34] implied no natural Type-S system can recognize its Hilbert consistency, and Willard subsequently [38,43,44] hybridized their formalisms with some techniques of Adamowicz-Zbierski [1,2] to establish that most Type-M systems cannot recognize their own semantic tableaux consistency.

Other fascinating efforts to evade the Second Incompleteness Theorem have used the Kreisel-Takeuti "CFA" system [22] or the the *interpretational framework* of Friedman, Nelson, Pudlák and Visser [10,24,27,33]. These systems are unrelated to our approach because they do not use Kleene-like *"I am consistent"* axiom-sentences. Instead, CFA uses the special properties of "second order" generalizations of Gentzen's *cut-free* Sequent Calculus, and the interpretational approach formalizes how some systems recognize their Herbrand consistency on localized sets of integers, which unbeknown to themselves, includes all integers. (These alternate results are interesting but unrelated to our approach.)

3 Defining Notation and Earlier Results

A function F will be called **Non-Growth** iff $F(a_1 ... a_j) \leq Maximum(a_1 ... a_j)$ holds. Six examples of non-growth functions are *Integer Subtraction* (where $x - y$ is defined to equal zero when $x \leq y$), *Integer Division* (where $x \div y$ equals x when $y = 0$, and it equals $\lfloor x/y \rfloor$ otherwise), $Maximum(x, y)$, $Logarithm(x)$ $Root(x, y) = \lceil x^{1/y} \rceil$ and $Count(x, j)$ designating the number of "1" bits among x's rightmost j bits. The term **U-Grounding Function** referred in [40] to a set of primitives, which included the preceding functions plus the *growth operations* of addition and $Double(x) = x + x$. Our language L^* was built out of these symbols, plus the primitives of "0", "1", "$=$" and "$\leq$".

In a context where t is any term in [40]'s language L^*, the quantifiers in the wffs $\forall\, v \leq t \;\; \Psi(v)$ and $\exists\, v \leq t \;\; \Psi(v)$ were called *bounded quantifiers.* Any formula in L^*, all of whose quantifiers are bounded, was called a Δ_0^* formula. The Π_n^* and Σ_n^* formulae were then defined by the usual rules that:

1. Every Δ_0^* formula is considered to be " Π_0^* " and also " Σ_0^* ".
2. A wff is called Π_n^* when it is encoded as $\forall v_1 \; ... \; \forall v_k \; \Phi$ with Φ being Σ_{n-1}^*
3. Also, a wff is called Σ_n^* when it is encoded as $\exists v_1 .. \exists v_k \; \Phi$, where Φ is Π_{n-1}^*.

Our article [40] used the symbol D to denote a deduction method. In addition to using Fitting's version of semantic tableaux methodology [8], it defined an alternative, called **Tab-k deduction**, that consisted of a speeded-up version of a tableaux, which permitted a *limited analog* of Gentzen-style deductive cuts for Π_k^* and Σ_k^* formulae.

Thus, if H denotes a sequence of ordered pairs (t_1, p_1), (t_2, p_2), ... (t_n, p_n), where p_i is a Semantic Tableaux proof of the theorem t_i, then H was called a **"Tab-k Proof"** of a theorem T from α's axioms iff $T = t_n$ and also:

1. Each of the "intermediately derived theorems" $t_1, t_2, \, ... \, , t_{n-1}$ have a complexity no greater than that of either a Π_k^* or Σ_k^* sentence.
2. Each axiom in p_i's proof either comes from α or is one of $t_1, t_2, \, ... \, , t_{i-1}$.

Let us say an axiom system α has a **Level-J Understanding** of its own consistency under a deduction method D iff α can prove that there exists no proofs using its axioms and D's deduction of both a Π_J^* theorem and its negation. In this notation, items A and B summarize [36, 38–40, 43]'s main results:

A. For any axiom system A using L^*'s U-Grounding language, [40] showed its $\text{IS}_D(A)$ formalism could prove all A's Π_1^* theorems and simultaneously verify its Level-1 consistency under Tab-1 deduction.

B. Two negative results, tightly complementing item A's positive result, were exhibited in [36, 38, 39, 43]. The first was that [36, 38, 43] showed most systems are unable to verify their Level-0 consistency under semantic tableaux deduction, when they included statement (3)'s "Type-M" axiom that multiplication is a total function. Moreover, [39] offered an alternate form of this incompleteness result, showing statement (2)'s *far weaker* Type-A systems cannot verify their Level-0 consistency under Tab-2 deduction.

The contrast between these positive and negative results had led to our conjecture that automated theorem provers are likely to eventually achieve a fragmentary part of the ambitions that were suggested by Hilbert in $**$. This is because the question of whether a formalism can support an *idealized Utopian* conception of its own consistency is *different* from exploring the degrees to which theorem-provers can possess a *fragmentary knowledge* of their own consistency. The Incompleteness Theorem has demonstrated an Utopian idealized form of self-justification is unobtainable, but our research has found some diluted cousins of this construct that are feasible

4 The $\mathrm{IS}_D(A)$ Axiom System

In a context where A denotes any axiom system using L^*'s U-Grounding language, $\mathrm{IS}_D(A)$ was defined in [40] to be an axiomatic formalism capable of recognizing all of A's Π_1^* theorems and corroborating its own Level-1 consistency under D's deductive method. It consisted of the following four groups of axioms:

Group-Zero: Two of the Group-zero axioms will define the constant-symbols, $\bar{c}_0$ and $\bar{c}_1$, designating the integers of 0 and 1. The Group-zero axioms will also define the growth functions of addition and $Double(x) = x + x$. The net effect of these axioms will be to set up a machinery to define any integer $n \geq 2$ using fewer than $3 \cdot \lceil \mathrm{Log}\; n \rceil$ logic symbols.

Group-1: This axiom group will consist of a finite set of Π_1^* sentences, denoted as F, which can prove any Δ_0^* sentence that holds true under the standard model of the natural numbers. (Any finite set of Π_1^* sentences F with this property may be used to define Group-1, as [40] noted.)

Group-2: Let $\ulcorner \Phi \urcorner$ denote Φ's Gödel Number, and $\mathrm{HilbPrf}_A(\ulcorner \Phi \urcorner, p)$ denote a Δ_0^* formula indicating p is a Hilbert-styled proof of theorem Φ from axiom system A. For each Π_1^* sentence Φ, the Group-2 schema will contain an axiom of form (4). (Thus $\mathrm{IS}_D(A)$ can trivially prove all A's Π_1^* theorems.)

$$\forall\, p \quad \{\ \mathrm{HilbPrf}_A(\ulcorner \Phi \urcorner, p) \quad \Rightarrow \quad \Phi\ \} \tag{4}$$

Group-3: The final part of the $\mathrm{IS}_D(A)$ will be a self-referencing Π_1^* axiom, indicating $\mathrm{IS}_D(A)$ meets §3's criteria of being "Level-1 consistent" under deductive method D. It is, thus, the following declaration:

> \# *No two proofs exist for a Π_1^* sentence and its negation, when D's deductive method is applied to an axiom system, consisting of the union of Groups 0, 1 and 2 with* **this sentence** *(looking at itself).*

One encoding of #, as a self-referencing Π_1^* axiom, appears in [40]. Thus, (5) is a Π_1^* styled encoding for # when: 1) $\mathrm{Prf}_{\mathrm{IS}_D(A)}(a,b)$ is a Δ_0^* formula indicating that b is a proof of a theorem a under $\mathrm{IS}_D(A)$'s axiom system and D's deduction method, and 2) $\mathrm{Pair}(x,y)$ is a Δ_0^* formula indicating that x is a Π_1^* sentence and that y represents x's negation.

$$\forall\, x\, \forall\, y\, \forall\, p\, \forall\, q \quad \neg\ [\ \mathrm{Pair}(x,y)\ \wedge\ \mathrm{Prf}_{\,\mathrm{IS}_{D(A)}}(x,p)\ \wedge\ \mathrm{Prf}_{\,\mathrm{IS}_{D(A)}}(y,q)\ ] \tag{5}$$

Notation. An operation $I(\ \bullet\ )$ that maps an initial axiom system A onto an alternate system $I(A)$ will be called **Consistency Preserving** iff $I(A)$ is consistent whenever all of A's axioms hold true under the standard model of the natural numbers. In this context, [40] demonstrated:

Theorem 1. *Suppose the symbol D denotes either semantic tableaux deduction or its Tab−1 generalization. Then the $IS_D(\ \bullet\ )$ mapping operation is consistency preserving (e.g. $IS_D(A)$ will be consistent whenever all of A's axioms hold true under the standard model of the natural numbers).*

We emphasize the most difficult part of [40]'s result was neither the definition of its $\mathrm{IS}_D(A)$'s axiom system nor the Π_1^* fixed-point encoding of (5)'s Group-3 axiom. Instead, the key challenge was the confirming of Theorem 1's "Consistency Preservation" property.

The confirming of this property is subtle because its invariant breaks down when D is a deduction method only slightly stronger than either semantic tableaux or Tab-1 deduction. Thus, Pudlák's and Solovay's work [27,31] implies Theorem 1's analog fails when D represents Hilbert deduction, and [39] showed its generalization fails even when D represents Tab-2 deduction.

5 A Finitized Generalization of Theorem 1's Methodology

One difficulty with $\mathrm{IS}_D(A)$ is that it employs an infinite number of different incarnations of sentence (4) in its Group-2 scheme (since it contains one incarnation of this sentence for each Π_1^* sentence Φ in L^*'s language). Such a Group-2 schema is awkward because it simulates A's Π_1^* knowledge almost via a brute-force enumeration.

Our Definition 1 and Theorems 2 and 3 will show how to mostly overcome this problem by compressing the infinite number of instances of sentence (4) in $\mathrm{IS}_D(A)$'s Group-2 schema into a purely finite structure.

Definition 1. Let β denote any finite set of axioms that have Π_1^* encodings. Then $\mathrm{IS}_D^{\#}(\beta)$ will denote an axiom system, similar to $\mathrm{IS}_D(A)$, except its Group-2 scheme will employ β's set of axioms, instead of using an infinite number of applications of statement (4)'s scheme. (Thus, the *"I am consistent"* statement in $\mathrm{IS}_D^{\#}(\beta)$'s Group-3 axiom will be the same as before, except that the *"I am"* fragment of its self-referencing statement will reflect these changes in Group-2 in the obvious manner.)

Theorem 2. *Let D again denote either semantic tableaux or Tab-1 deduction, and β again denote a set of Π_1^* axioms. Then $IS_D^{\#}(\beta)$ will be consistent whenever all β's axioms hold true under the standard model. (In other words, $IS_D^{\#}(\beta)$ will satisfy an analog of Theorem 1's consistency preservation property for $IS_D(A)$.)*

Theorem 2's proof is almost identical to [40]'s proof of Theorem 1. It will. not be repeated in this extended abstract. Instead, this section will apply Theorem 2 to show how **finite-sized** self-justifying logics can provide an **infinite amount** of "kernelized" Π_1^* information.

Definition 2. Let $\mathrm{Test}_i(t,x)$ denote any Δ_0^* formula, and $\ulcorner \Psi \urcorner$ denote Ψ's Gödel number. Then $\mathrm{Test}_i(t,x)$ will be called a **Kernelized Formula** iff Peano Arithmetic can prove every Π_1^* sentence Ψ satisfies (6)'s identity:

$$\Psi \quad \Longleftrightarrow \quad \forall\, x \;\; \mathrm{Test}_i(\, \ulcorner \Psi \urcorner \,,\; x \,) \tag{6}$$

There are infinitely many Δ_0^* predicates $\mathrm{Test}_1(t,x)$, $\mathrm{Test}_2(t,x)$, $\mathrm{Test}_3(t,x)$ ... satisfying this kernelized condition (one of which is illustrated by Example 1). An enumerated list of all the available kernels is called a **Kernel-List**.

Example 1. The set of true Σ_1^* sentences is r.e. This implies there exists a Δ_0^* formula, called say Probe(g, x), such that g is the Gödel number of a Σ_1^* statement that holds true in the Standard Model iff (7) is true:

$$\exists\, x \quad \text{Probe}(g, x) \;\wedge\; x \geq g \tag{7}$$

Now, let Pair(t, g) denote a Δ_0^* formula that specifies t is the Gödel number of a Π_1^* statement and g is the Σ_1^* formula which is its negation. Then our notation implies that t is a true Π_1^* statement if and only if (8) holds true:

$$\forall\, x \quad \neg\, [\, \exists\, g \;\leq\; x \quad \text{Pair}(t, g) \;\wedge\; \text{Probe}(g, x) \;\,] \tag{8}$$

Thus if Test$_0(t, x)$ denotes the Δ_0^* formula of $\neg\,[\,\exists\, g \;\leq\; x$ Pair$(t, g) \wedge$ Probe$(g, x)]$, it is one example of what Definition 2 would call a "Kernelized Formula".

Definition 3. Let us recall Definition 2 defined **Kernel-List** to be an enumeration of all the kernelized formulae Test$_1(t, x)$, Test$_2(t, x)$, Test$_3(t, x)$... . Assuming Test$_i(t, x)$ is the $i-$th element in this list and Ψ is an arbitrary Π_1^* sentence, Ψ's **i-th Kernel Image** will be defined as the following Π_1^* sentence:

$$\forall\, x \quad \text{Test}_{\,i}\,(\, \ulcorner\, \Psi\, \urcorner\, ,\; x\;) \tag{9}$$

Example 2. The Definitions 2 and 3 suggest that there is a subtle relationship between a sentence Ψ and its $i-$th Kernel Image. This is because Definition 2 indicates that Peano Arithmetic can prove the invariant (6), indicating that Ψ is equivalent to its $i-$th Kernel Image. However, a weak axiom system can be plausibly uncertain about whether this equivalence holds.

Thus if a weak axiom system proves statement (9) (rather than Ψ), it may not be able to equate these results. This problem will apply to Theorem 3's formalism. However, Theorem 3 will be still of much interest because §6 will illustrate a methodology that overcomes many of Theorem 3's limitations.

Theorem 3. *Let A denote any system, whose axioms hold true in arithmetic's standard model, and i denote the index of any of Definition 2's kernelized formulae Test$_i(t, x)$. Then it is possible to construct a finite-sized collection of Π_1^* sentences, called say $\beta_{A,i}$, where $IS_D^{\#}(\beta_{A,i})$ satisfies the following invariant:*

> *If Ψ is one of the Π_1^* theorems of A then $IS_D^{\#}(\beta_{A,i})$ can prove (9)'s statement (e.g. it will prove the "the $i-$th kernelized image" of Ψ).*

Proof Sketch: Our justification of Theorem 3 will use the following notation:

1. Check(t) will denote a Δ_0^* formula that produces a Boolean value of "True" when t represents the Gödel number of a Π_1^* sentence.
2. HilbPrf$_A\,(t, q)$ will denote a Δ_0^* formula that indicates q is a Hilbert-style proof of the theorem t from axiom system A .
3. For any kernelized Test$_i(t, x)$ formula, GlobSim$_i$ will denote (10)'s Π_1^* sentence. (It will be called A's $i-$th **"Global Simulation Sentence"**.)

$$\forall\, t\ \ \forall\, q\ \ \forall\, x\ \ \{\ \ [\ \ \text{HilbPrf}_A\,(t,q)\ \ \wedge\ \ \text{Check}(t)\ \ ]\ \ \Longrightarrow\ \ \text{Test}_i(t,x)\ \ \}\tag{10}$$

In this notation, the requirements of Theorem 3 will be satisfied by any version of the axiom system $\text{IS}_D^{\#}(\beta)$, whose Group-2 schema β is a finite sized consistent set of Π_1^* sentences that has (10) as an axiom. (This includes the minimal sized such system, that has only (10) as an axiom.) This is because if Ψ is any Π_1^* theorem of A, whose proof is denoted as $\bar{p}$, then both the Δ_0^* predicates of $\text{HilbPrf}_A\,(\ulcorner\Psi\urcorner,\bar{p})$ and $\text{Check}(\ulcorner\Psi\urcorner)$ are true. Moreover, $\text{IS}_D^{\#}(\beta)$'s Group-1 axiom subgroup was defined so that it can automatically prove all Δ_0^* sentences that are true. Thus, $\text{IS}_D^{\#}(\beta)$ will prove these two statements and hence corroborate (via axiom (10)) the further statement:

$$\forall\, x\ \ \text{Test}_{\,i}\,(\ \ulcorner\Psi\urcorner\ ,\ x\ )\tag{11}$$

Hence for each of the infinite number of Π_1^* theorems that A proves, the above defined formalism will prove a matching statement that corresponds to the $i-$th kernelized image of each such proven theorem. □

6 L-Fold Generalizations of Theorem 3

Theorem 3 is of interest because every axiom system A will have its formalism $\text{IS}_D^{\#}(\beta_{A,i})$ prove the $i-$th kernelized image of every Π_1^* theorem that A proves. This fact is helpful because (6)'s invariance holds for all Π_1^* sentences. Moreover, our "U-Grounded" Π_1^* sentences capture all Conventional Arithmetic's *crucial* Π_1 information because they can view multiplication as a 3-way Δ_0^* predicate $\text{Mult}(x,y,z)$ via (12)'s encoding of this predicate.

$$[\,(x=0\vee y=0)\Rightarrow z=0\,]\ \wedge\ [\,(x\neq 0\wedge y\neq 0\,)\ \Rightarrow\ (\ \frac{z}{x}=y\wedge\ \frac{z-1}{x}<y\ \ )\,]\tag{12}$$

One difficulty with $\text{IS}_D^{\#}(\beta)$ and $\text{IS}_D^{\#}(\beta_{A,i})$ was mentioned by Example 2. It was that while Peano Arithmetic can corroborate (6)'s invariance for every Π_1^* sentence Ψ , these latter systems cannot also do so.

While there will probably never be a perfect method for fully resolving this challenge, there is a pragmatic engineering-style solution that is often available. This is essentially because our proof of Theorem 3 employed a formalism β that used essentially only one axiom sentence (e.g. (10)'s Π_1^* declaration).

Since the $\text{IS}_D^{\#}(\beta)$ formalism was intended for use by any finite-sized system β, it is clearly possible to include any finite number of formally true Π_1^* sentences in β. Thus for some fixed constant L, one can easily let β include L copies of (10)'s axiom framework for a finite number of different Test_1, Test_2 ... Test_L predicates, each of which satisfy Definition 2's criteria for being kernelized formulae. In this case, $\text{IS}_D^{\#}(\beta)$ will formally map each initial Π_1^* theorem Ψ of some axiom system A onto L resulting different Π_1^* theorems of the form (9).

Remark 1. Our basic conjecture is, essentially, that a goodly number of issues, concerning logic-based engineering applications called say E, may have convenient solutions via self-justifying logics, that follow the preceding outlined L-fold strategy. Thus, we are suggesting that if β is a large-but-finite set of axioms, that consists of L copies of (10)'s axiom framework for different $\text{Test}_1 ... \text{Test}_L$ predicates, then some future engineering applications E may possibly have their needs met by an $\text{IS}_D^{\#}(\beta)$ formalisms, when a software engineer meticulously chooses an appropriately constructed finite-sized β.

Remark 2. The preceding was not meant to overlook that the Second Incompleteness Theorem is a robust result, applying to all logics of sufficient strength. Our suggestion, however, is that computers are becoming so powerful, in both speed and memory size as the 21st century is progressing, that there will likely emerge engineering-style applications E that will benefit from $\text{IS}_D^{\#}(\beta)$'s self-referencing formalisms when a *large-but-finite-sized* β is delicately chosen. Moreover, it is of interest to speculate whether such computers can partially imitate a human being's approximate instinctive conjectures about his own consistency (that, as common colloquially held conjectures, seem to serve as *essential prerequisites* for humans to gain their motivation to cogitate).

Sections 7-9 will examine the preceding issues in further detail. One of their themes will be that our exceptions to Gödel's second theorem, while sometimes nontrivial, clearly do not narrow the main intentions of Gödel's result.

7 Comparing Type-M and Type-A Formalisms

Let us recall axioms (1)-(3) indicated Type-A systems differ from Type−M formalisms by treating Multiplication as a 3-way relation (rather than as a total function). For the sake of accurately characterizing what our systems can and cannot do, we have described our results as being fringe-like exceptions to the Second Incompleteness Theorem, from the perspective of an Utopian view of Mathematics, while perhaps being more significant results from an engineering-style perspective of knowledge. Our goal in this section will be to amplify upon this perspective by taking a closer look at Type-A and Type-M formalisms.

Let us assume that $x_0 = 2 = y_0$ and that $x_1, x_2, x_3, ...$ and $y_1, y_2, y_3, ...$ are defined by the recurrence rules of:

$$x_{i+1} = x_i + x_i \qquad \text{AND} \qquad y_{i+1} = y_i * y_i \tag{13}$$

The sequences $x_0, x_1, x_2, ...$ and $y_0, y_1, y_2, ...$ will thus represent the growth rates associated with the addition and multiplication primitives, lying in the statements (2) and (3)'s **"Type-A"** and **"Type-M"** axioms.

Since $x_0 = 2 = y_0$, the rule (13) implies $y_n = 2^{2^n}$ and $x_n = 2^{n+1}$. The $y_0, y_1, y_2, ...$ sequence will, thus, grow much more quickly than the $x_0, x_1, x_2, ...$ sequence (since y_n's binary encoding will have an $\text{Log}(y_n) = 2^n$ length while x_n's binary encoding will have a shorter length $= \text{Log}(x_n) = n + 1$).

Our prior papers noted that the difference between these growth rates was the reason that [36, 38, 43] showed all natural Type-M systems, recognizing integer-multiplication as a total function, were unable to recognize their tableaux-styled consistency — while [35, 37, 40] showed some Type-A systems could simultaneously prove all Peano Arithmetic Π_1^* theorems and corroborate their own tableaux consistency. Their gist was that a Gödel-like diagonalization argument, which causes an axiom system to become inconsistent as soon as it proves a theorem affirming its own tableaux consistency, stems, ultimately, from the exponential growth in the series $y_0, y_1, y_2, ...$.

This growth, thus, facilitates an intense amount of self-referencing, using the identity $\text{Log}(y_n) \cong 2^n$, that will, ultimately, invoke the force of Gödel's seminal diagonalization machinery. It thus raises raises the following question:

> $***$ How natural are exponentially growing sequences, such as $y_0, y_1, y_2..$, whose $n-$th member needs 2^n bits for its encoding, when such lengths are greater than the number of atoms in the universe when merely $n > 100$? Is such a sequence's use, for corroborating the Second Incompleteness Effect, an inherently artificial construct ?

We will not attempt to derive a Yes-or-No answer to Question $***$ because it is one of those epistemological questions that can be debated endlessly. Our point is that $***$ probably does not require a definitive positive or negative answer because both perspectives are useful. Thus, the theoretical existence of a sequence integers of $y_0, y_1, y_2, ...$, whose binary encodings are doubling in length, is tempting from the perspective of an Utopian view of mathematics, while awkward from an engineering styled perspective. We therefore ask: *"Why not be tolerant of both perspectives? "*

One virtue of this tolerance is it ushers in a greater understanding for the statements $*$ and $**$ that Gödel and Hilbert made during 1926 and 1931. This is because the Incompleteness Theorem demonstrates no formalism can display an understanding of its own consistency in an idealized Utopian sense. On the other hand, §6 suggested these remarks by Gödel and Hilbert might receive more sympathetic interpretations, if one sought to explore such questions from a less ambitious almost engineering-style perspective.

Our main thesis is supported by a theorem from [42]. It indicated that tableaux variations of self-justifying systems have no difficulty in recognizing that an infinitized generalization of a computer's floating point multiplication (with rounding) is a total function. The latter differs from integer-multiplication, by not having its output become double the length of its input when a number is multiplied by itself. Thus, the intuitive reason [42]'s multiplication-with-rounding operation is compatible with self-justification is because it avoids the inexorable exponential growth under rule (13)'s sequence $y_0, y_1, y_2..$.

Also, Theorem 4 indicates self-justifying logics can view a double-precision form of integer multiplication as likewise a total function. Its proof, exactly analogous to [42]'s methodology, will appear in a longer version of this paper.

Theorem 4. *Let us assume the A in $IS_D(A)$ and $IS_D^{\#}(\beta_{A,i})$ represents Peano Arithmetic. Then $IS_D(A)$ and $IS_D^{\#}(\beta_{A,i})$ can formalize two total functions, called Left(a,b) and Right(a,b), where any pair of integers (a,b) is mapped onto the left and right halves of a and b's multiplicative product.*

Remark 3. One slightly tricky aspect is that our positive results, involving [42]'s floating point multiplication and Theorem 4's double precision multiplication, *should not be confused* with a different examination of Herbrandized consistency in [44]. The latter took advantage of the fact that our Herbrand-styled proofs, in [44]'s paradigm, are exponentially longer than their tableaux counterparts, thus allowing [44] to formalize a limited use of multiplication (because its deductive methods was exponentially less efficient). Thus [44]'s results, while perhaps theoretically interesting, are basically irrelevant to engineering environments, e.g. the main concern of Theorems 1–4 (especially in regards to their particular interpretations given in Remark 2).

8 A Different Type of Evidence Supporting Our Thesis

Let us recall Pudlák and Solovay [27, 31] observed that essentially all Type-S systems, containing merely statement (1)'s axiom that successor is a total function, cannot verify their own consistency under Hilbert deduction. (See also related work by Buss-Ignjatovic [5], Švejdar [32] and the Appendix A of [37])

It turns out that [39] generalized these results to show that Equation (2)'s Type-A systems are unable to verify their own consistency, under the Tab-2 deduction (defined in §3). At the same time, the IS_D and $\mathrm{IS}_D^{\#}$ frameworks, from Sections 4 and 5, can verify their own consistency under Tab-1 deduction. Our goal in this section will be to illustrate how the tight contrast between these positive and negative results is analogous to the differing growth rates of the sequences $x_0, x_1, x_2, ...$ and $y_0, y_1, y_2, ...$ from rule (13).

During our discussion $G_i(v)$ will denote the scalar-multiplication operation that maps an integer v onto $2^{2^i} \cdot v$. Also, Υ_i will denote the statement, in the U-Grounding language, that declares that G_i is a total function. Our paper [39] proved that Υ_i has a Π_2^* encoding. It is also implied that G_i satisfied:

$$G_{i+1}(v) \quad = \quad G_i(\ G_i(v)\ ) \tag{14}$$

It was noted in [39] that this identity implies one can construct an axiom system β, comprised of solely Π_1^* sentences, where a semantic tableaux proof can establish Υ_{i+1} from $\beta + \Upsilon_i$ in a constant number of steps. This implies, in turn, that a Tab-2 proof from β will require no more that $\mathrm{O}(n)$ steps to prove Υ_n (when it uses the obvious n-step process to confirm in chronological order Υ_1 , Υ_2 , ...Υ_n .)

These observations are significant because $G_n(1) = 2^{2^n}$. Thus, [39] showed a Tab-2 proof from β can verify in $\mathrm{O}(n)$ steps that this integer exists.

This example is helpful because it illustrates the difference between the growth speeds under Tab-1 and Tab-2 deduction, is analogous to the differing growth rates of the sequences $x_0, x_1, x_2, ...$ and $y_0, y_1, y_2, ...$ from rule (13). Hence once again, a faster growth-rate will usher in the Second Incompleteness Theorem's power (e.g. see [39]).

This analogy suggests that the Second Incompleteness Theorem has different implications from the perspectives of Utopian and engineering theories about the intended applications of mathematics. Thus, a Utopian may possibly be comfortable with a perspective, that contemplates sequences $y_0, y_1, y_2, ...$ with elements growing in length at an exponential speed, but many engineers may be suspicious of such growths.

A hard-core engineer, in contrast, might surmise that the inability of self-justifying formalisms to be compatible with Tab-2 deduction is not as disturbing as it might initially appear to be. This is because Tab-2 differs from Tab-1 deduction by producing exponential growths that are so sharp that their material realization has no analog in the everyday mechanical reality that is the focus of an engineer's interest.

Our personal preference is for a perspective lying half-way between that of an Utopian mathematician and a hard-nosed engineer. Its dualistic approach suggests some form of diluted partial agreement with Hilbert's goals in $**$.

9 Related Reflection Principles

An added point is that there are many types of self-justifying systems available, with some better suited for engineering environments than others.

Ideally, one would like to develop self-justifying systems S that could corroborate the validity of (15)'s reflection principle for all sentences Φ.

$$\forall p \ \ [\ Prf_S^D(\ulcorner \Phi \urcorner, p) \ \ \Rightarrow \ \ \Phi \ ] \tag{15}$$

Löb's Theorem establishes, however, that all systems S, containing Peano Arithmetic's strength, are able to prove (15)'s invariant *only in the degenerate case* where they prove Φ itself. Also, the Theorem 7.2 from [37] showed essentially all axiom systems, *weaker* than Peano Arithmetic, are unable to prove (15) for all Π_1^* sentences Φ simultaneously. Thus, Theorem 5 will be near optimal:

Theorem 5. *For any input axiom system A, it is possible to extend the self-justifying $IS_D(A)$ and $IS_D^{\#}(\beta_{A,i})$ systems, from Theorems 1 and 3, so that the resulting self-justifying logics S can also:*

1. *Verify that Tab-1 deduction supports the following analog of (15)'s self-reflection principle under S for any Δ_0^* and Σ_1^* sentences Φ :*

$$\forall p \ \ [\ Prf_S^{\mathrm{Tab}-1}(\ulcorner \Phi \urcorner, p) \ \ \Rightarrow \ \ \Phi \ ] \tag{16}$$

2. *Verify (17)'s more general* **"root-diluted"** *reflection principle for S whenever θ is Σ_1^* and Φ is a Π_2^* sentence of the form "$\forall u_1 ... \forall u_n \ \ \theta(u_1 ... u_n)$ ".*

$$\forall p \ [\ Prf_S^{\mathrm{Tab}-1}(\ulcorner \Phi \urcorner, p) \ \Longrightarrow \ \forall x \ \forall u_1 < \sqrt{x} \ \ ... \ \ \forall u_n < \sqrt{x} \ \ \theta(u_1 ... u_n) \ ] \tag{17}$$

Theorem 5's proof will rest upon hybridizing the techniques from [37]'s tangibility reflection principle with Theorem 3's methodologies, in a natural manner, as will be demonstrated in a longer version of this article. Analogous to our other results, Theorem 5 reinforces the theme about how exceptions to the Second Incompleteness Theorem may appear to be *quite minor* from the perspective of an Utopian view of mathematics, while being significant from an engineering standpoint. In Theorem 5's particular case, this is because:

A. The ability of Theorem 5's system S to support (16)'s self-reflection principle under Tab−1 proofs for any Δ_0^* and Σ_1^* sentence, as well as to support (17)'s root reflection principle for Π_2^* sentences, is clearly significant.

B. The incompleteness result of [37]'s Theorem 7.2 imposes, however, sharp limitations upon Item A's generality (in that it cannot be extended to fully all Π_1^* sentences, *in an undiluted sense).*

Thus, the tight fit between A and B is reminiscent of other slender borderlines, that separated generalizations and boundary-case exceptions for the Incompleteness Theorem, explored earlier. Once again, the Second Incompleteness Theorem is seen as robust, from an idealized Utopian perspective on mathematics, while permitting caveats from engineering styled perspectives.

This dualistic viewpoint allows one to nicely share *partial (and not full)* agreement with Hilbert's main aspirations in $**$, while also appreciating the stunning achievement of the Second Incompleteness Theorem.

10 Concluding Remarks

At a purely technical level, this article has reached beyond our prior papers in several respects, including §5's demonstration that any initial system A can have a kernelized image of its Π_1^* knowledge duplicated by $\mathrm{IS}_D^{\#}(\beta_{A,i})$'s **strictly finite sized** self-justifying system, and also by Section 6's and Remark 2's quite pragmatic L-fold generalizations of this result.

These results help resolve the mystery that has enshrouded the Second Incompleteness Theorem and the statements $*$ and $**$ of Gödel and Hilbert. This is because we have *meticulously separated* the goals of a pristine theoretical study of mathematical logic from those of a *finite-sized* axiomatic subset of mathematics, intended for modeling mostly an engineering environment.

There is no question that Gödel's Second Theorem is ideally robust, relative to a purely pristine approach to mathematics. On the other hand, we suspect Hilbert was *half-way correct* by speculating in $**$ about humans possessing a knowledge about their own consistency, *in at least some strikingly weak and tender sense,* as essentially a fundamental prerequisite for *psychologically motivating* their cogitations.

Thus in a context where the limitations of axiom systems, that fail to recognize multiplication as a total function, are manifestly obvious, even when such systems duplicate Peano Arithmetic's central Π_1^* knowledge, it is legitimate to inquire whether posterity might find some *partial-albeit-not-full* redeeming value in formalisms having *weak-style* knowledges of their Tab−1 consistency?

More precisely, Sections 5-9 were intended to provide a unified broad-scale interpretation of our diverse earlier results from [35–44]. In a context where the Incompleteness Theorem is firmly understood to be sufficiently ubiquitous to preclude Hilbert's aspirations in $**$ from ever being fully realized, they show how some *fragmentary portion* of Hilbert's conjectures can be corroborated by *judiciously weakened* logics, using a formalism, that is *much less* than ideally robust, *although not fully immaterial.*

Acknowledgments. I thank Bradley Armour-Garb and Seth Chaiken for many useful suggestions about how to improve the presentation of our results. I also thank the anonymous referees for their comments. This research was partially supported by NSF Grant CCR 0956495.

References

1. Adamowicz, Z.: Herbrand consistency and bounded arithmetic. Fundamenta Mathematicae 171(3), 279–292 (2002)
2. Adamowicz, Z., Zbierski, P.: On Herbrand consistency in weak theories. Archive for Mathematical Logic 40(6), 399–413 (2001)
3. Bezboruah, A., Shepherdson, J.C.: Gödel's second incompleteness theorem for Q. Jour. Symb. Logic 41(2), 503–512 (1976)
4. Buss, S.R.: Bounded Arithmetic. Studies in Proof Theory, Lecture Notes 3. Published by Bibliopolis (1986) (Revised version of Ph. D. Thesis)
5. Buss, S.R., Ignjatovic, A.: Unprovability of consistency statements in fragments of bounded arithmetic. Annals Pure and Applied Logic 74(3), 221–244 (1995)
6. Dawson, J.W.: Logical Dilemmas the life and work of Kurt Gödel. AKPeters (1997)
7. Feferman, S.: Arithmetization of metamathematics in a general setting. Fundamenta Mathematicae 49, 35–92 (1960)
8. Fitting, M.: First Order Logic and Automated Theorem Proving. Springer (1996)
9. Friedman, H.M.: On the consistency, completeness and correctness problems. Technical report, Ohio State Univ. (1979), See Pudlák [28]'s summary of this result
10. Friedman, H.M.: Translatability and relative consistency. Technical report, Ohio State Univ. (1979), See Pudlák [28]'s summary of this result
11. Gödel, K.: Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. Monatshefte für Math. Phys. 38, 173–198 (1931)
12. Gödel, K.: The present situation in the foundations of mathematics. In: Feferman, S., Dawson, J.W., Goldfarb, W., Parson, C., Solovay, R. (eds.) Collected Works Volume III: Unpublished Essays and Lectures, pp. 45–53. Oxford University Press (2004), Our quotes from this 1933 lecture come from its page 52
13. Goldstein, R.: Incompleteness The Proof and Paradox of Kurt Gödel. Norton (2005)
14. Hájek, P.: Mathematical fuzzy logic and natural numbers. Fundamenta Mathematicae 81, 155–163 (2007)
15. Hájek, P.: Towards metamathematics of weak arithmetics over fuzzy. Logic Journal of the IPL 19(3), 467–475 (2011)

16. Hájek, P., Pudlák, P.: Metamathematics of First Order Arithmetic. Springer (1991)
17. Hilbert, D.: Über das Unendliche. Mathematische Annalen 95, 161–191 (1926)
18. Hilbert, D., Bernays, P.: Grundlagen der Mathematik, vol. II. Springer (1939)
19. Jeroslow, R.G.: Consistency statements in formal theories. Fundamenta Mathematicae 72, 17–40 (1971)
20. Kleene, S.C.: On notation for ordinal numbers. Jour. Symb. Logic 3(1), 150–156 (1938)
21. Kołodziejczyk, L.A.: On the Herbrand notion of consistency for finitely axiomatizable fragments of bounded arithmetic theories. Jour. Symb. Logic 71(2), 624–638 (2006)
22. Kreisel, G., Takeuti, G.: Formally self-referential propositions in cut-free classical analysis and related systems. Dissertationes Mathematicae 118, 1–50 (1974)
23. Löb, M.H.: A solution of a problem of Leon Henkin. Jour. Symb. Logic 20(2), 115–118 (1955)
24. Nelson, E.: Predicative Arithmetic. Math. Notes. Princeton Univ. Press (1986)
25. Parikh, R.: Existence and feasibility in arithmetic. Jour. Symb. Logic 36(3), 494–508 (1971)
26. Parsons, C.H.: On $n-$quantifier induction. Jour. Symb. Logic 37(3), 466–482 (1972)
27. Pudlák, P.: Cuts, consistency statements and interpretations. Jour. Symb. Logic 50(2), 423–442 (1985)
28. Pudlák, P.: On the lengths of proofs of consistency. In: Collegium Logicum: Annals of the Kurt Gödel, vol. 2, pp. 65–86. Springer-Wien (1996)
29. Rogers, H.A.: Theory of Recursive Functions and Effective Compatibility. McGraw Hill (1967)
30. Salehi, S.: Herbrand consistency of some arithmetical theories. Jour. Symb. Logic 77(3), 807–827 (2012)
31. Solovay, R.M.: Telephone conversation in 1994 describing Solovay's generalization of one of Pudlák's theorems [27], using some methods of Nelson and Wilkie-Paris [24, 34] (The Appendix A of [37] offers a 4-page summary of this conversation) (1994)
32. Svejdar, V.: An interpretation of Robinson arithmetic in its Grzegorczjk's weaker variant. Fundamenta Mathematicae 81, 347–354 (2007)
33. Visser, A.: Faith and falsity. Annals Pure and Applied Logic 131(1), 103–131 (2005)
34. Wilkie, A.J., Paris, J.B.: On the scheme of induction for bounded arithmetic. Annals Pure and Applied Logic 35, 261–302 (1987)
35. Willard, D.E.: Self-verifying axiom systems. In: Mundici, D., Gottlob, G., Leitsch, A. (eds.) KGC 1993. LNCS, vol. 713, pp. 325–336. Springer, Heidelberg (1993)
36. Willard, D.E.: The semantic tableaux version of the second incompleteness theorem extends almost to Robinson's arithmetic Q. In: Dyckhoff, R. (ed.) TABLEAUX 2000. LNCS, vol. 1847, pp. 415–430. Springer, Heidelberg (2000)
37. Willard, D.E.: Self-verifying systems, the incompleteness theorem and the tangibility reflection principle. Jour. Symb. Logic 66(2), 536–596 (2001)
38. Willard, D.E.: How to extend the semantic tableaux and cut-free versions of the second incompleteness theorem almost to Robinson's arithmetic Q. Jour. Symb. Logic 67(1), 465–496 (2002)

39. Willard, D.E.: A version of the second incompleteness theorem for axiom systems that recognize addition but not multiplication as a total function. In: Hendricks, V., Neuhaus, F., Pederson, S.A., Scheffler, U., Wansing, H. (eds.) First Order Logic Revisited, pp. 337–368. Logos Verlag, Berlin (2004)
40. Willard, D.E.: An exploration of the partial respects in which an axiom system recognizing solely addition as a total function can verify its own consistency. Jour. Symb. Logic 70(4), 1171–1209 (2005)
41. Willard, D.E.: A generalization of the second incompleteness theorem and some exceptions to it. Annals Pure and Applied Logic 141(3), 472–496 (2006)
42. Willard, D.E.: On the available partial respects in which an axiomatization for real valued arithmetic can recognize its consistency. Jour. Symb. Logic 71(4), 1189–1199 (2006)
43. Willard, D.E.: Passive induction and a solution to a Paris-Wilkie open question. Annals Pure and Applied Logic 146(2), 124–149 (2007)
44. Willard, D.E.: Some specially formulated axiomizations for $I\Sigma_0$ manage to evade the Herbrandized version of the second incompleteness theorem. Information and Computation 207(10), 1078–1093 (2009)
45. Yourgrau, P.: A World Without Time: The Forgotten Legacy of Gödel and Einstein. Basic Books (2005), See page 58 for the passages we have quoted