Recording of Scientific Notes Notarized by Dan E. Willard on May10, 2014

Our proof of $P \neq NP$ will certainly be non-trivial. It has, thus, taken us 21+ years to discover how [37]'s initial 1993 formalism for generating quite *innocent-looking* boundary-case forms of exceptions to the Second Incompleteness Theorem can be revised to formulate a proof of $P \neq NP$. A nice aspect of this proof is that it is much *easier* to appreciate with good retrospective hindsight than in foresight (where one begins with a *full blind start* that will inevitably cause many false turns to be made). This is because there are multiple misleading mirages, that lie along the roadway towards proving $P \neq NP$, that can be avoided, with good healthy retrospective advice.

Our formal proof of $P \neq NP$ will appear in Sections ???-???. The next three sections will explain how our core methodologies are related to some seminal observations by Pudlák, Solovay, Nelson and Wilkie-Paris [27, 32, 23, 35], along with some related work by Buss-Ignjatovic and Švejdar [6, 33].

Their results establish that logical axiom systems, employing Hilbert-style deductive methods, obey more powerful generalizations of the Second Incompleteness Theorem than the traditional classic generalizations of Gödel's result, developped by Hilbert-Bernays and Löb [15, 22]. The appreciation of the importance of the seminal observations made by Pudlák, Solovay, Nelson and Wilkie-Paris, was one of the vital steps that has led to proving $P \neq NP$.

These notes were finished on May 1, 2014, but not notarized until May10 because my Wollic paper was accepted on May 2 (and rewriting it required my attention).

NEW *Section's Title Background Notation*

Throughout this article, a will denote a recursively enumerable set of "proper axioms", and d will denote a deduction method, used for deriving theorems from a. Such d will typically include some rules of inferences (such as modus ponens) accompanied by some "logical axioms", denoted as L_d, that we consider to be part of d's infrastructure rather than a's infrastructure. (This seemingly minor notational point, which views L_d as part of the deduction method d, meticulously seperated from the set of proper axioms a, shall greatly simplify several aspects of our proof. We therefore emphasize that by the axiom system " a ", our discourse will always be referring to a set of proper axioms that is, automatically, further extended by L_d's logical axioms.)

Example 0.1 The deduction methods d, that we will use, are quite conventional. They can, for instance, correspond to the paradigms that were used in the textbooks by way of Enderton, Mendelson or Papadimitriou [7, 7, 7]. Thus, d can represent:

1. Enderton's 6-part set of logical axioms, augmented by solely modus ponens as a rule of inference, when one uses [7] as a framework to define L_d. (This framework was also employed in Papadimitriou's textbook [7]).
2. Mendelson's framework [7], which has two inference rules of modus ponens and generalisation, and which would have L_d include the five pure logical axioms of A1-A5 (from section 2.3 of [7]) plus the two equality axioms A6 & A7 (defined in its Section 2.8).

In general, we will use the term Hilbert Deduction to refer to a deductive methodology that includes a modus ponens rule and whose efficiency differs from the above two proof methodologies by no more than a polynomial increase in proof length. Our proof of $P \neq NP$ can employ essentially any Hilbert-style deductive method. We will focus our attention mainly on Item 1 and 2's paradigms for the sake of simplicity because they are probably best known to a wide audience of readers. (The best source about the more advanced properties of Hilbert deduction can probably be found in the Hájek- Pudlák textbook [13], but its more advanced techniques will not be used during our proof of $P \neq NP$.)

Remark 0.2 Some deduction methods d fall into a non-Hilbert category, where they possess the same logical power as Example 0.1's "Hilbert" methods, while they are super-exponentially less efficient in the worst case. This category includes the semantic tableaux and resolution methods, used in automated theorem proving [7]. Despite their worst-case disadvantage, these techniques produce, often, good heuristics in automated deduction. (For instance, [37, 40, 45, 46] illustrate how tableaux-style methods can roughly approximate how human beings assess a limited (but real) understanding of their own consistency, that helps explain how humans motivate themselves to cogitate.) These results will interest logicians, interested in the broader capacities of self-justifying logics, but they will not be germane to our proof of $P \neq NP$. For simplicity, the latter proof will employ *solely* the better known Hilbert-style methods, defined in Example 0.1.

REMOVE Thus happily, this second lengthy group of articles, published by us, can be fully ignored by a reader who wishes to focus mostly on our proof of $P \neq NP$.)

Definition 0.3 Once again, let a denote an axiom system, and d denote a deduction method. The ordered pair (a, d) will be called **Self Justifying** when:

i one of a's theorems (or at least one of its axioms) will state that the deduction method d, applied to the system a, will produce a consistent set of theorems, and

ii the axiom system a is in fact consistent.

Example 0.4 For any (a, d), it is easy to construct a system $a^d \supseteq a$ that satisfies the Part-i condition. For instance, a^d could consist of all of a's axioms plus an added "SelfRef(a, d)" sentence, defined as stating:

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

Kleene [17] discussed how to encode approximate analogs of this "SelfRef(a, d)" statement. Each of Kleene, Rogers and Jeroslow [17, 29, 16] noted a^d may, however, be inconsistent (despite SelfRef(a, d)'s assertion), thus causing it to violate Part-ii of self-justification's definition. This is because if the ordered pair (a, d) is too strong then a classic *Gödel-style diagonalization* argument can be applied to the axiom system $a^d = a + \mathit{SelfRef}(a, d)$, where the added presence of the statement SelfRef(a, d) causes this extended version of a to ironically become automatically inconsistent. Thus, the machinery of the sentence "SelfRef(a, d)", while relatively easy to encode via an application of the Fixed Point Theorem, is ironically *not often useful*.

We are now ready to explain the intuition that has motivated our 21+ year quest to arrive at a proof of $P \neq NP$, via the employment of self-justifying axiom systems. It was that the final three italicized words in the preceding paragraph(e.g. the phrase *"most often useless"*) mean nothing other than exactly what they suggest. Thus, our goal has been to examine, in meticulous detail, the tiny class of boundary-case formalisms that evade the typical reach of Gödel's Second Incompleteness Theorem, with the objective of finding a means to apply some one of these formalisms to demonstrate $P \neq NP$.

For the sake of clarity, our proof of $P \neq NP$ will ultimately be a theorem proven by a *fully classical* logic, whose correctness is beyond doubt. The reasons for our interest in nonclassical logics, that evade the Second Incompleteness Theorem, is that these strange-looking entities will be a useful intermediate step to examine during our proof of $P \neq NP$. Hence, our proof of $P \neq NP$ will be arrived by exploring how fully classical (and uncontroversial) logics view unusual self-justifying logics as a simplifying helpful intermediate step.

NEW *Section's Title The Combined Work of Pudlák, Solovay, Nelson and Wilkie-Paris*

Throughout this article, $Add(x, y, z)$ and $Mult(x, y, z)$ will denote two 3-way predicate symbols specifying $x + y = z$ and $x \cdot y = z$. An axiom system a will be said to recognize successor, addition and multiplication as Total Functions iff it includes 1-3 as axioms.

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

Also, an axiom system a will be called Type-M iff it contains (1) – (3) as axioms, Type-A iff it contains only (1) and (2) as axioms, and Type-S iff it contains only (1) as an axiom. Moreover, a will be called Type-NS iff it contains none of these axioms.

Our initial paper [37] about self-justification involved Type-A logics that can formalize their own consistency under semantic tableaux deduction. Shortly after we published this result, Robert Solovay telephoned us, in April of 1994, indicating he knew how to strengthen a variant of the Second Incompleteness Theorem due to Pavel Pudlák [27], with additional techniques developed by Nelson and Wilkie-Paris [35, 23], to show that no reasonable Type-S axiom system (that treats Addition and Multiplication as 3-*way* relations) can verify its own Hilbert consistency. This result, summarized by Theorem 2.1, improves upon Bezboruah-Shepherdson's earlier work in [3].

> **Theorem 3.1** (*Solovay's 1994 modification* [32] *of Pudlák's 1985 formalism* [27] *using the added methodologies of Nelson and Wilkie-Paris* [23, 35]) *Let a denote any consistent Type-S system that can verify addition and multiplication satisfy their usual associative, communitative, distributive and identity axioms. Then a cannot prove a theorem affirming its own consistency under any of Example 0.1's forms of Hilbert-style deduction.*

We will never use Theorem 3.1 during our proof of $P \neq NP$. Its importance in achieving this result cannot, however, be understated. This is because Theorem 3.1 *indicates* that only Type-NS arithmetics are plausibly capable of verifying their own Hilbert consistency.

It thus indicates any hope in using self-justifying logics to prove $P \neq NP$ must center around the daunting properties of astonishingly weak Type-NS systems. (I have no doubt that my 21+ *years effort to prove* $P \neq NP$ would have taken more than roughly — giggle, giggle — say 210+ years — if I was not aware of Theorem 3.1's crucial negative result and the tiny class of Type-NS boundary-case exceptions to the Second Incompleteness Theorem that it *implicitly* permits.)

SOME RELEVANT BACKGOUND INFORMATION One certainly eclectic aspect of Theorem 3.1 is that there is currently no available published proof of its statement. Essentially what had transpired was I published in 1993 a proof that relatively rich Type-A logics can verify their own consistency under semantic tableaux and resolution forms of deduction [37]. Robert Solovay subsequently telephoned me in April of 1994 [32] to tell me that he knew how to generalize a theorem of Pudlák to show that a variant of Theorem 3.1 would pertain to Hilbert deduction.

Solovay was reaching the age where Berkeley would grant him a very lucrative early retirement in 1994. He showed no interest in writing up either Theorem 3.1's proof or spending more that five minutes discussing its subject matter (together with sending me a 10-sentence email, outlining the core aspects of his fascinating methodology) and suggesting I read the less-general results by Pudlák, Nelson and Wilkie-Paris [27, 23, 35].)

In a 2009 blog, Richard Lipton [21] has indicated that Solovay was always reluctant to write up the proofs of his results. For instance, this blog reports Solovay was the only mathematician to ever remove tenure from Berkeley's Math Department before he published a single paper. Also, many logicians [8, 13, 18, 28, 27, 28, 33, 34, 35] have commented about Solovay's unusual reluctance to publish many results he derived, analogous to Theorem 3.1, that logicians have greatly valued.

Thus at the *suggestion of an anonymous referee*, our first JSL article [50] included a deliberately abbreviated 4-page Appendix A that formalized a slightly weaker but-more-easily comprehensible variant of Theorem 3.1. Other insightful results, *partially* but not *fully analogous* to Theorem 3.1, were published by Buss-Ignjatovic, Švejdar and ourselves in [5, 33, 43]. Also, Pudlák's initial paper [27] comes very close to establishing Theorem 3.1's result. Furthermore, Pudlák specifies in [28] that some *also-unpublished* work by Harvey Friedman [9] runs in some partially but-not-fully analogous directions to Pudlák's work on proof lengths.

In any case, the reader will be happily relieved to learn our proof of $P \neq NP$ will never use the formalism of the mysterious Theorem 3.1, lying in an analog of the tales "Neverland" from a "Peter Pan" novel by *J. M. Barrie*, as an intermediate step. Instead, the sole function of this never-published theorem is to serve as a signal about what steps a proof of $P \neq NP$ must delicately avoid.

This is because there are no examples of viable Type-S logics that verify their Hilbert style consistency. Our proof of $P \neq NP$ will thus use Theorem 3.1 as a beacon, signaling that only the annoyingly weak Type-NS formalisms can plausibly support a proof of $P \neq NP$, via self-justifying logics. (Beyond this useful guiding hint, our proof of $P \neq NP$ will not use Theorem 3.1's special machinery.)

NEW *Section's Title Getting Started*

Our first two papers about self-justifying axiom systems were the conference papers [37, 38] whose unconventional results were thankfully readily accepted by the respective teams of editors at Gottlob, Leitsch & Mundici and Beame & Buss. The former paper [37] established that relatively strong axiom systems that satisfied Equation (2)'s Type-A requirement could verify their own semantic tableaux consistency. The latter article [38] indicated some Type-NS systems could corroborate their own Hilbert consistency and speculated whether a proof of $P \neq NP$ could be established via a souped-up version of these results.

The Sections 7-10 of [38] briefly outlined the approximate reasons for [38]'s core conjecture. The 18 years that had separated the years 1996 from 2014 obviously indicates that much additional work was needed to complete this task.

5-10-14

Jodi Sherman Schultz
No. 01SH6181014
Notary Public, State of New York
Qualified in Albany County
Commission Expires 01/22/20 16

It turns out that there are many different variants of Type-NS arithmetics that can can verify different forms of their own Hilbert consistency. It was not until the year 2006 that we would introduce in [47] two versions of Type-NS self-verifying arithmetics whose hybridisation would ultimately lead to the proof of $P \neq NP$. The remainder of this section will roughly summarise [47]'s core results. Our earlier 1993 and 1996 announcements in [37, 38] are of historical interest, but only [47]'s more mature year-2006 formalism will be central for proving $P \neq NP$.

During our discoure, the term "Introspective Semantics" will refer to the analog of our self-justifying axiom systems from Definition 0.3 and Example 0.4 that use analogs of SelfRef(α, d)'s fixed-point axiomatic sentence to enable an axiom system to look at itself and declare that *"I am consistent"*. It is, convenient to have such "Introspective Semantic" systems have their names always begin with the acronym "IS", as a useful reminder.

In a context where β is an initial axiom system, that contains at least the logical power of Peano Arithmetic (PA), the two main variants of intospective systems from [47] carried the names of "ISCE(β)" and "ISINF(β)". Both were Type-NS self-justifying formalisms that could recognise their own Hilbert consistency and also prove all the Π_1 theorems that β could prove under a slightly modified language where Add(x, y, z) and Mult(x, y, z) represent 3-way predicates symbols that formalise addition and multiplication. The remainder of this section will offer a capsulised summary of ISCE(β)'s and ISINF(β)'s properties and discuss [47]'s related generalisation of the Second Incompleteness Theorem, which demonstrates that these two formalisms are near-maximal.

The formalism ISCE(β) shall avoid using Equation (1)'s Type-S axiom sentence, declaring successor is a total function, by instead employing an infinite number of built-in constant symbols $C_0, C_1, C_2, C_3 \ldots$ for defining the set of positive integers. The constant symbols C_0 and C_1 will represent the integers of 0 and 1. Each other C_j will be defined to represent the quantity 2^{j-1} . These integers will be defined via essentially an "Additive Naming Convention" indicating $C_{j+1} = C_j + C_j$. Since ISCE(β) techaically does not contain an additive function symbol, its definition of C_{j+1} will formally rest upon using Equation (4)'s 3-way addition predicate symbol:

$$\text{Add}\,(C_j\,,\,C_j\,,\,C_{j+1}) \tag{4}$$

We will assume the "name" of the built-in constant symbol "C_j" is encoded using $O[\,\text{Log}(j+2)\,]$ bits. The advantage of using an "additive naming convention", which assigns names to only integers which are powers of 2, is that its methodology will nicely assure that the powers of 2 have integer names that are shorter than the lengths of their binary encodings.

Since the ISCE(β) system will contain a built-in function symbol for representing integer-subtraction as a total function (where $x - y$ is defined to be equal to zero when $x < y$), Equation (4)'s additive naming convention can clearly define any integer that is not a power of 2. For instance since $10 = 16 - 4 - 2$, the integer 10 can be represented as "$C_4 - C_2 - C_1$".

A detailed definition of ISCE(β) is provided in [47]. It uses an analog of Example 0.4's "SelfRef" axiomatic sentence to corroborate its own consistency. The Theorem 3 of [47] indicated ISCE(β) is a self-justifying formalism that verifies its own Hilbert consistency. In essence, ISCE(β) evades the Pudlák-Solovay variant of the Second Incompleteness Theorem because it is a Type-NS system. At the same time, the Theorem 4 from [47] demonstrated that if one used an alternate naming convention, where say C^*_j equals $2^{2^{j-2}}$ when $j \geq 2$, then the force of the Second Incompleteness Theorem would return, *even though the said formalism is a Type-NS arithmetic.* In particular, [47]'s generalisation of the Second Incompleteness Theorem will apply to settings where one replaces Equation (4)'s additive naming convention with (5)'s "Multiplicative Naming Convention" :

$$\text{Mult}\,(C^*_j\,,\,C^*_j\,,\,C^*_{j+1}) \tag{5}$$

It turns out that if a reader wishes to glance at our article [47], then he can afford to *entirely omit* its Section 5 and Theorems 4 & 5. This is because the latter, unlike [47]'s Theorem 3, are unrelated to the proof of $P \neq NP$. Indeed, I would recommend that readers, interested in mainly NP's characterisation, intially omit [47]'s Section 5 and its Theorems 4 & 5 because the latter involve a complicated proof that is ultimately unrelated to NP's fundamental properties.

A peculiar aspect of [47] is that it does contain one other result, that we recently discovered to be crucial for proving $P \neq NP$, although we previously presumed Theorem 6 was too specialised for it to have much significance. Its awkward-but-useful formalism involves the following definition:

Definition 0.5 Let us assume that α is an axiom system that contains a Predecessor function symbol, where Pred(x) = Min($x - 1, 0$), as well as contains a constant symbol C_1 for representing the value "1". Also, let Pred^j(x) denote a functional operation that consists of j iterations of such a predecessor function (e.g. this notation implies that k is the unique integer that satisfies the identity $\text{Pred}^{k-1}(k) = C_1$). Then α will be said to possess Infinte Far Reach iff there exists some *finite subset* of α's set of proper axioms, called say γ, such that for every integer k the finite system γ is capable of proving (6)'s invariant (that intuitively states the integer quantity k does exist).

$$\exists x \quad \text{Pred}^{k-1}(x) = C_1 \tag{6}$$

The Theorem 6 of [47] shows it is possible to construct awkward-but-viable self-justifying arithmetics with infinite far reach that are capable of corroborating their own Hilbert consistency and proving the validity of all of Peano Arithmetics Π_1 theorems, using again a slightly modified language that treats addition and multiplications as 3-way relations (rather than as function primitives). In particular, [47]'s ISINF formalism can achieve this property without violating the Pudlák-Solovay version of the Second Incompleteness Theorem because it is a Type-NS formalism, formally incapable of verifying any of the operations of Successor, Addition or Multiplication are total functions. Moreover, this "ISINF" class of formalisms has a fundamentally different anatomy from [47]'s alternate "ISCE" variant of self-justifying logics because the latter uses an *infinite number* of different instances of the axiom schema (4) to construct the full infinite range of integers.

Thus while many aspect of [47] "ISINF" style formalism are awkward and super-impractical (due to the unwieldy long proofs it produces), this framework differs from the alternative ISCE mechanism by, at least, showing that self-justifying axiom systems with Infinite Far Reach are theoretically capable of confirming their own Hilbert consistency.

The intuition behind our proof of $P \neq NP$ is that we noticed that if this invariant was false then a hybridisation of the ISCE and ISINF frameworks will produce a contradiction (showing that the invariant $P = NP$ cannot possibly hold). In particular assuming that there is available an algorithm a that can solve length$-n$ SAT problems on a Turing machine in say n^b time, one can apply the ordered pair (a, b) to develop two types of self-justifying axiom systems, called "Is.Bulky" and "Is.Super.Bulky", that are natural hybrids of the ISCE and ISINF frameworks, with the following pleasing combination of properties:

I. A natural generalisation of the techniques, used to prove [47]'s Theorems 3 and 6, will imply that both Is.Bulky and Is.Super.Bulky are efficent and consistent, *when one uses* the ordered pair (a,b) to construct the axiomatic input β for our Is.Bulky and Is.Super.Bulky frameworks.

II. In contrast to Item 1's positive result, a generalisation of Gödel-style diagonalisaton argument will imply both of Is.Bulky and Is.Super.Bulky are inconsistent (essentially because they are too efficient to escape the reach of the Second Incompleteness Theorem).

Our proof of $P \neq NP$ will essentially rest on the fact that Items I and II are incompatible with each other whenever an ordered pair (a, b) formally represents a polynomial solution for SAT problems. (We will, thus, derive the result $P \neq NP$, via a proof by contradiction, that shows an ordered pair (a, b) would otherwise produce two incompatible results.

Cheerful News about Misleading Mirages: An earlier section of this article indicated that there existed several types of mirages, which would tend to make most computer researchers and logicians overlook the essence of the *proof of $P \neq NP$.* Some cheerful news, that will be now revealed, is that our proof of $P \neq NP$ becomes much easier to conceptualise, intuitively, when one is aware of two mirages that can be initially quite misleading.

The first mirage concerns the question about whether Gödel's Second Incompleteness Theorem is 100 % rather than 99 % ubiquituous, when it describes the inability of formalisms to recognise their own consistency. A theme of our reseach in [37]–[49] is that the the Second Incompleteness Theorem is obviously 100 % technically correct, as well as 100 % germane to conventional mathematical paradigms. However, our research perspective has been that seemingly tiny 1 % categories of boundary-case exceptions can be helpful in solving many open questions, including $P \neq NP$.

REMOVE However, we view the tiny 1 % category of boundary-case exceptions as also revealing. In a sense the same passion to examine that *final seemingly miniscule* tiny remaining 1 % category of possibilities, that led to the Fredman-Willard Fusion Trees at the 1990 STOC and FOCS conferences [?, ?], will now lead to a proof of $P \neq NP$.

The second surprising mirage also caught this author off guard. It concerns [47]'s "ISINF" axiomatic framework, which our APAL-2006 paper had called "awkward" and "artificial". These adjectives accurately characterise "ISINF". One might even, actually, harshly add that ISINF *is so fully inefficient and purely artificial* in its defining structure that one *can confidently predict* THAT IT IS INCAPABLE *of ever producing any unequivocally positive result.*

Such italicised and bold-faced words, which I certainly did not dare insert into [47]'s published text, are, technically, accurate. Yet in April of 2014, I discovered, that *they were also a mirage!* This is because the statement "$P \neq NP$" is a *negative assertion*, discussing the *inability* of polynomial time algorithms to simulate a non-deterministic Turing machine. The latter class *of negativist* statements have quite different anatomies and logical *properties* than *positive-styled* theorems!

Thus although ISINF's anatomy is unlikely to produce any useful *strictly positive* results, it is a useful vehicle in a proof by contradiction, which shows that a hybridisation of the ISINF and ISCE frameworks makes it impossible for $P = NP$ to hold (e.g. see Items I & II on the prior page).

(via the 2p-art argument that was roughly summarised by the Items I and II, earlier in this section).
NEW Section's Title Added Notation and Further Intuitions
imply the combination of Item 1's efficiency results, together with prohibitions imposed by the Incompleteness Theorem will imply that either and both of the
either $P \neq NP$ or any other fundamental fact.
dd42.3

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

Old abstract *This article will confirm the validity of Cook's seminal $P \neq NP$ conjecture, a topic whose foundational implications have been discussed also by Karp and Levin. We had conjectured at a 1996 Dimacs Workshop, co-chaired by Sam Buss and Paul Beame, that a proof of $P \neq NP$ could be obtained by somehow hybridising the mechanics of Gödel Diagonalization Arguments with those of the unusual boundary-case exceptions pertaining to Gödel's Second Incompleteness Theorem, that we have called "Self-Justifying Logics".*

The above conjecture was, of course, highly counter-intuitive because Gödel's First Incompleteness Theorem (which asserts the inherent undecidablity of conventional arithmetic) and his Second Incompleteness Theorem (which establishes the inability of conventional logics to confirm their own consistency) are both known to be amazingly robust and resilient twin results. However after researching this topic during the last essentially 21+ years (and publishing six papers about it in the Journal of Symbolic and Annals of Pure and Applied Logic), we have now, finally, developed a method to demonstrate P $\neq$ NP by means of a proof-by-contradition that exploits the inherent tension and divergent properties that separate the seemingly miniscule-but-valid boundary-case exceptions for Gödel's Second Incompleteness Theorem from sophisiticated generalizations of Gödel-style diagonalization arguments.

Our discourse will presume that the reader has a sufficient knowledge about Mathematical Logic to understand the main themes in at least one of the graduate-level textbooks by say Enderton, Mendelson or Papadimitriou [?, ?, ?].

NEW Section's Title NEED

Throughout this article, α will denote an axiom system, and d will denote a deduction method. An ordered pair (α, d) 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 system $\alpha^d \supseteq \alpha$ that satisfies the Part-i condition. 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 axiom system α with this sentence "$SelfRef(\alpha, d)$" (looking at itself).

Kleene [17] discussed how to encode approximate analogs of this "$SelfRef(\alpha, d)$" statement. Each of Kleene, Rogers and Jeroslow [17, 29, 16] noted α^d may, however, be inconsistent (despite $SelfRef(\alpha, d)$'s assertion), thus causing it to violate Part-ii of self-justification's definition.

This problem arises in settings more general than Gödel's paradigm, where α was an extension of Peano Arithmetic. There are many settings where the Second Incompleteness Theorem does generalize [1, 2, 3, 4, 5, 6, 8, 10, 11, 12, 13, 15, 19, 20, 22, 24, 27, 28, 29, ?, 32, 33, ?, 35, 41, 43, 48]. Each such result formalizes a paradigm where self-justification is infeasible, due to a diagonalization issue. Many logicians have, thus, hesitated to employ a $SelfRef(\alpha, d)$ axiom because $\alpha^d = \alpha + SelfRef(\alpha, d)$ is typically inconsistent. [1]

Our research explored special circumstances [40, 45, 46, 47] where it is feasible to construct self-justifying formalisms. These paradigms involved weakening the properties a system can prove about addition and/or multiplication (to avoid the preceding difficulties). To be more precise, let $Add(x, y, z)$ and $Mult(x, y, z)$ denote two 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 1-3 as axioms.

$$\forall x \; \exists z \quad Add(x, 1, z) \tag{7}$$
$$\forall x \; \forall y \; \exists z \quad Add(x, y, z) \tag{8}$$
$$\forall x \; \forall y \; \exists z \quad Mult(x, y, z) \tag{9}$$

A logic system α will be called **Type-M** iff it contains (1) – (3) as axioms, **Type-A** iff it contains only (1) and (2) as axioms, and **Type-S** iff it contains only (1) as an axiom. Also a system is called **Type-NS** iff it contains none of these axioms. The significance of these constructs 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 [46]. Also, [47] 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 [23, 27, 32, 36] implied no natural Type-S system can recognize its Hilbert consistency, and Willard subsequently [41, 48, 49] 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.

An interesting aspect of (a) and (b) is that there is a tight match between their positive and negative results from a quantitative perspective, but they do not reach one's idealized goals for deduction. This is because there is something about the deep yearnings of deductive reasoning, roughly suggested by Hilbert in $**$ and Gödel in $*$, which has not been addressed. This remaining gap will be explored by our Theorems ??, ??, ?? and ??.

Gödel's Incompleteness Theorem is a 2-part result. Its first half indicates no decision procedure can identify all the true statements of arithmetic. Its Second Incompleteness Theorem specifies sufficiently strong systems cannot verify their own consistency. Gödel was careful to insert the following caveat into his historic paper [10], indicating a diluted form of Hilbert's Consistency Program could be successful:

> $*$: "It must be expressly noted that Proposition XI (e.g. the Second Incompleteness Theorem) 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 M or in A)."

Yourgrau has summarized, in detail, Gödel's considerations about this subject. Thus [38] indicated that "for several years" after [10]'s publication, Gödel "was cautious not to prejudge" whether some unusual formalism, different from Peano Arithmetic, might provide some type of proof of its own consistency.

OLDER NOTES 18 NOTES

On Tuesday March 19 while going to Dr. Seltzer's office, I corrected my old proof of the Pudlak conjecture to make it work. I also consider it possible, albeit not likely, that it could lead to a proof of Cook's conjecture. I can prove Pudlak's conjecture both for Atjai-style pigeon-hole growth functions and their generalization for scan functions that maps powers of constant c onto other powers of c. (Indeed, it should work for numbers of the form $2^{P(i)}$ where $P(i)$ is a polynomial of i.) This finishes my old paper about Pudlak's conjecture about generalized Atjai operation, which I will hencefoth call "opaque" fuctions. A function that full defines all values of $F(i)$ will be a "**completely transparent**" function.

It is best to prove Pudlak's conjecture by using a modified form of semantic tableau, called X-tableau, that allows for all instances of the Law of Excluded Middle to be used as axioms. All other Frege-style proofs can be reduced to this formalism with no more than a Polynomial growth in proof length, that should actually be linear in most cases. For instace, this applies to a deduction method, called EM, which is idential to Enderton's deduction method except that his Group-1 axioms are replaced by Mendelson first three logical axiom schema. Also we will call a proof a **Normalized Tableaux** proof relative to an axiom Φ if Φ appears in the root of such a X-tableaux tree and nowhere else.

Only for the sake of notational simplicity, we will assume our semantic trees are so normalized and that they have no growth function symbols(e.g. they represent a function $F(x_1, x_2, ... x_j)$ as a j+1 way relation $R_f((x_1, x_2, ... x_j, y)$ where y is the function's output)

Then all the results of my old paper at the GWU Math conference will niele generalize. I think my new method will called something like **Judicious Blind Probing**. If such a X-tab "**candidate tree**" has n nodes then we will say it has an L-model "M" with $L = n + 2$ when M is a model of the natural numbers that assumes that all integers $\leq L + 1$ exists and our Opague Atjai-like function F satisfies the following rules

$$\forall x \; F(x) \neq 0 \tag{10}$$

$$\forall x \leq L \quad F(x) \leq L + 1 \tag{11}$$

$$\forall x \, \forall y \quad [illegible] \tag{12}$$

Moreover, we will say a node v in the candidate tree is L-Consistent in a Φ-**blind sense** iff there exists some L-Model M where the sentence v and all its ancestors with the possible exception of the tree root Φ satisfy the model M under an interpretation where each of the named constants C in this proof tree assume values $\leq L$. The magic bullet, used in the proof is the observation that if G is the godel number of the of the minimal proof of absurdity in our self verifying axiom system, whose "I am consistent declaration" is stored in its root called Φ then at least one top down tree branch ending at node v will be L-Consistent in a Φ-**blind sense** when L is a number less than say $\frac{G}{2}$. (Indeed with a little extra work that I will not do now, this proof will work when L equals $Log(G)$.) The point is that I know how to make this methodology produce my needed proof by contradiction.

Moreover, I can generalize my mehtod in many directions. One issues is that it is probably best for my paper to concentrate on the simplest version of Atjai function that are analgous to the incremental naming convention, and then to proceed to the additive naming convention at the end of the article (which obviously will also support my method when one walks through powers of 2 because i have enough wiggle room available). Other issues are listed below

1. Method generalizes when a proof is defined not by one integer p but rather a k-tuple of such integers
2. If k=2 this implies that my method will know that there exists no two simulatneous proof of a sentence and its negation (both encoded in prenex normal form for simplicity).
3. Also, it will apply to many versions of the tangibility reflection princible, including the TangRoot princible and stronger

Somewhere I will wish to use the words from 3 days ago of "blind specimen as starting with a permutation p of L integers, blind case, blind specimens, Starting Blind Set, and make Reduced Blind SET.

NOTES ADDED MARCH 23 afternoon (starting at 1.30pm after exercising for 75 minutes at YMCA) I saw Nicolson on Friday at 11am. and she told me that her guess was that a co-author would make the writing of the paper harder (rather than easier. (Therefore, I will not ask AG for help in writing the paper, but I might ask him to read it and thank him in the Acknowledgement, similar to my prior papers.) Also, I went over the basic idea on Sunday morning and I convinced myself they were correct. (Interestingly, I could not reconstruct my prior idea without first reading my March 21 notes (below) and then filling in the missing pieces. This is partly because I took a 2-day rest from my new project and returned to the old Wollic project!) Another new point is that I decided to call my core idea **Opaque Pseudo-Model Analysis (OPMA)**. It rests on the idea of taking an X-tableaux proof whose only axiom you doubt is a Π_1^- sentence called Φ and making Φ the root of the proof tree (appearing nowhere else. Then you entertain the possibility that Φ is false in the Standard Model. In that case you can create a model consisting of all α's axioms except that Φ is replaced by $\neg \Phi$. The call of OPMA analysis is to show that the existence of the resulting model is contradictory, which is why I "judiciouley" call it a pseudo-model. The trick is to use the combination of the false presumption of the existence of the pseudo-model M to construct a path p through the proof tree where all nodes, except "its root" Φ hold in the model M. That is impossible because the proof tree has each node end with a contradiction. (The contradiction essentially arises because the bottom nodes of the proof tree, being Δ_0^- sentences stemming from Φ must be correct. This is because if Φ is sentence of the form $\forall x \;\; \theta(x)$ then all constants symbols c constructed by the proof are small enough to have $\theta(c)$ be true when we assume our proof tree T is **MINIMUM CONTRADICTORY PROOF**.) I wonder if this technique is related to Forcing in Cohen's discussion of the Axiom of Choice?

NOTES ADDED APRIL 11, 2014: I am traveling tomorrow to Washington DC to join my son Robert and celebrate my wife's birthday. I am certain my new methods will show that if P=NP then there will exist a transparent polylog system of infinite far reach, and they likely can also prove $P \neq NP$ itself. The diagonalization sentence, needed to do this is the statement: "No short proof of there exists a short proof of this sentence" (This sentence will be called $**$.)

By short proof p we mean a proof p having the property that a temporarily-hypothesized algorithm a for solving SAT problems it time bn^b cannot find in p units of time existence of a proof of 0=1 from our self-referencing axiom system and ALSO having the minor additional caveats (from essentially my dimacs paper) that essentially $LogLog(p)$ is greater than some trivially specified constant K.

I will use the term ISWISH(a,b,A) to refer temporarily to my analog of the old ISINP(A) system in this short note. Its name will be later presumably changed in the future to perhaps "Is.Problematic" or better-yet "Is.Troubling" or "Is.Vexing". It will be used to prove $P \neq NP$ by us showing that it displays contradictory behaviour when (a,b) is a generic method for solving SAT problems in polynomial time. It will differ from ISINP(A) in the following repects:

A) It is harmless and makes matters simpler for ISWISH(a,b,A) to include the axioms of the multipliative naming convention.

B) The Group-3 axiom of IsWish will include the self-reflecting statement that on each occassion when it is proven that our (a,b) methodology in time t is able to demonstate the non-existence of a proof of $0 = 1$ of length less than L, the constant $\sqrt{2}^L$ exists.

C) It is possible that IsWISH's group-3 will also include an analog of ISINP's statement "There is no proof of $0 = 1$ from me, but I am not sure if that statement is necessary, but I am 99 % certain that it is necessary!

Interestingly I have been going in circles during the last week until I realised the above triple diagonalization sentence in $**$ seems to work, while the alternate version below (that has one extra use of the word "no" is folly:)

"No short proof of there exists NO short proof of this sentence" (e.g. second use of the word "no" should be avoided, and it has been causing me difficulties for many years).

The intuition behind our proof of $**$ will be to establish it by contradiction. Thus if $**$ was false then there would be a short proof P of the inside part of $**$ (which I will henceforth call $*$). This short proof of $*$ will demonstrate the existence of a short proof Q of $**$. Due to the shortness of P and Q, we can establish in one single computable proof that both

1) Q exists as a short proof, and

2) A copy of its proof implies the validity of $**$

But the point is the combination of (1) and (2) amounts to a proof of $0 = 1$ because they contradict each other. This leads to our desired proof by contradiction by showing that Q cannot exist (in formal contradiction to P's statement).

Or more precisely, I think (now as I meditate about it) that this last step of our proof-by-contractiction needs the fact that my proof leads to ISWISH(a,b,A) contradicting its "I am consistent" declaration. (This is why I added, in the last couple of minutes, a caveat to Item c that I was now, after 3-4 hours of doubt forcing me to eat an unhealthy shrimp roll, concluding with 99 % certainty that this part of Group-3 is indeed needed, as my intuition had always told me).

NOTES ADDED APRIL 18, 2014: wwww The theorem essentially held up during the last week of my meditations. I had some doubts about it after I did April 11 notary and thought it was wrong when driving with Irina to washinton D.C on april 12. I started the night of April 12 feeling discouraged, but then I woke up while from my sleep around 3am on April 13, realising the theorem worked and its proof was easier than I thought when the formalism was moderately revised. The idea is to use a vaariant of my ISWish formalism, almost the same as on April 11, whose name I will probably call Is.Problematic or Is.Troubling or Is.Challenging

Its definition will require using a SAT(a,b) formalism for the sake of contradiction, (defined on April 11). Its Group-3 will use this system to grow larger numbers, and it will use a variant of the additive naming convention that defines a series of finite constants $c_1, c_2 ... c_M$ where the VERY HELPFUL upper bound M depends on (a,b). (This bound was not in my April 11 notes.) It will also contain the new group-3 clause that states if a proof of 0=1 of length L has been found then no integer exists that is larger than say $Max(C_m, 2^{\sqrt{L}})$. (This clause makes it unnecessary for group-3 to include the sentence "There is no proof of 0=1 from me" as an axiom because it can make such into a theorem. However, the retaining this as an axiom is probably helpful in simplifying the proof.)

---

[1] Typically, $\alpha^d = \alpha + SelfRef(\alpha, d)$ will be inconsistent because a standard Gödel-like self-referencing construction will produce a proof of $0 = 1$ from α^d, even when α is consistent.

TERRY A. PE
Notary Public,
Reg. No
Qualifie
Commissi

The diagonalization sentence that I will use is simpler that its April 11 counterpart. Let us say a number N is J-Large iff say $LogLog(N) > J$. Then the diagonlization sentence can be as simple as the statement "There is no J-Large proof of me from the system Is.Troubling". I know that I used this variant of a Godel sentence before, and my vague memory is that analogs of it appeared in my Dimacs 1998 paper. (As an alternative one would have J-Large mean $Log^J N > 2$, and I am not sure which definition of J will make my paper easier to read.)

NOTES ADDED APRIL 2, shortly before midnight: aaaa. After having double yesterday, I now am certain the Rogue function theory will work. In our discussion an axiom system α that includes the sentence Ψ will be said on L, Ψ Mutated model iff the latter model assumes E is a set of integers $< L$ and V is the set of values associated with the constant symbols W that appear in a proof tree, where $V \subset E$. We will assume that our universal quantifiers have a range in V and the existential quantifiers have a range in E. Our use of law of excluded middle will involve sentences of the form $\phi \vee \neg\phi$ where both ϕ and $\neg\phi$ are written in prenex normal form. (This means both can be simultaneously satisfied because of the way E and W are defined.)

Our conclusion sentence Ψ will be sentence $\forall x \neg Rougue(x, 0)$ where $Rougue(x, y)$ is the relation describing our Rogue function. Our L, Ψ Mutated Model M will assume be a model of $\alpha - \Psi$ that assumes all integers belong to E and that for each constant symbol $c_i \in W$ there will be the additional constraints that $Val(C_i) \in V$ and also that $\neg Rougue(c_i, 0)$. The latter implies that a proof tree showing α is inconsistent will model-theoretically affirm the statement " $\neg Rougue(c_i, 0)$". We can then use a counting argument to show that at least one branch of the candidate tree will have all its statements holding true in our model where universal and existential quantifiers have the respective ranges of V and E. This will assure that candidate tree cannot be a contradiction proof, thus finishing our task. Our proof will assume that L equals the godel number of the smallest proof of $0=1$ from α and show that this is impossible because a smaller proof will be forced to exist under our model. (This is because our Rogue relationship is defined to be a function in the sense that $\forall x \exists y$ $Rogue(x, y)$ is true.)

NOTE ADDED AFTER MIDNIGHT TO ABOVE. I forget to add that our L, Ψ Mutated Model requires the existence of a 1-to-1 function $R(x)$ mapping integers $< L$ into the same where this mapping from E to E has the property that if $Val(C_i) = x$ then $R(x) \neq 0$ for the case of R simulates an Atjai pigeon-whole function. For the case of power-of-2 generalization, the comparable R is defined by

1) $R(x) = x$ when x is not power of 2.
2) $R(x)$ is a power of 2 when x is a power of 2, and if $Val(c_i) = x$ then $R(x) \neq 1$

Perhaps law of excluded middle should look like the expression:
Expression following was in beginquote $\forall x_1 \forall x_2 .. \forall x_y \phi(x_1, x_2 .. x_n) \vee \neg \phi(c_1, c_2 .. c_n)$ where both $\phi(x_1, x_2 .. x_n)$ and $\neg \phi(c_1, c_2 .. c_n)$ are prenex normalized.

NOTES ADDED MARCH 31, shortly before midnight: After having a week of doubts, I am now certain that the Rogue function method will work. The idea is to make the first line of semantic tableaux candidate tree to be the sentence "$\forall x \ \phi(x)$ where $\phi(x)$ says there is no proof of $0=1$ from me, and to have only the bottom nodes of the proof tree be statements of the form $\phi(c)$. Such a proof tree will be called "normalized". Also, if $R(x)$ is a rogue function that maps powers of 2 onto powers of 2 then our analog of an Atjai Rogue definition will state that "if g is power of 2 such that there exists no proof $p < g$ of the sentence of $0=1$ from me then $R(g) \neq 1$." Then one can use the counting argument that I constructed earlier this month to infer that no minimal proof of $0=1$ can exist. (An interesting sidenote is that I was confused over the last few days about whether my old argument would would work. Instead, I focused on proving that if $P \neq NP$ then an obnovious middle-level exception to the Second Incompleteness THeorem would exist by modifying the ISINF axioms system from my APAL 2006 paper. I was unhappy with that argument, however, because I suspected my revised version of the Atjai method would produce the same result. I figured out how to do this between 8.30 and 9.30 tonight, and then went to eat at My Father Place Bar before writing up these notes. This is a much stronger result than before.

In a future paper, I intend to use the term Is.Mini and Is.Middle to describe the miniaturized and middle-level "boundary-case exceptions" to the Second Incompleteness that arise when one uses respectively Pudalk's version of Atjai function and my power-of-2 walking generalization of it.

MINOR AMENDMENT DONE ON 11.45pm MARCH 31 to above notes Above was essentially except that the sentence "if g is power of 2 such that there exists no proof $p < g$ of the sentence of $0=1$ from me then $R(g) \neq 1$." should have "$p < g$" replaced with "$p << g$" where latter states that p's length must be less than say g's length+3

NOTES ADDED MARCH 21 Between 2 and 3am (after sending email to AC about our tomorrow telephone call): I decided to uses the terminology Rogue Function Description to describe my generalization of Atjai functions and Cantorized Rogue Functions to describe the usual generalization of Rogue functions where the base axiom system β allows for an uncountably infinite number of different allowed rogue functions to meet its definition. A conventional 20-th century non-Rogue function would be called an non-Rogue Function or perhaps better yet a Conventional non-Rogue-like Functional Object or equivalently Conventional non-Rogue-like Functional Definiton or perhaps a non-Rogue Standard-Form Functional Definiton abreviated as NRSF. my evasion of the Second Incompleteness Theorem obviously rests on using an axiom system of infinite far reach whose only Growth functions are Rogue fucntions. I am tentatively asking Brad to become a co-author (about one week from now) and for our paper to be written in a Socrates-Plato format where Brad's voice (the "Plato") summarizes Willard's ideas (the "Soctrates"). I will insist that I be the chronological first author of this paper. (That seems fair because because Brad has 30 years of life ahead of him, and I have high blood pressure and Diabetes.) I have also thought about the title of my paper. Some ideas are How the Use of Non-Conventional Rogue Functions Can Be Used to Develop Axiom Systems that Can "Appreciate" their Own Consistency in a Surprising Strong Sense While Simultaneously Respect Gödel's Prohibition Against Axiom Systems "Proving" their Own Consistency in More than A Weak Sense. After thinking about the issue during 4 to 8am, I am not yet sure whether I should write this article alone, as opposed to inviting Brad as a co-author. Also an alternate title would be: A Reply to a Question by Pavel Pudlak about the Implications of Atjai Pigeon-hole functins to Willard's Prior Work about Self-Justifying Axioms Systems aaaa END OF April 18, 2014 Notary

REST CAN BE TRIVIALLY DELETED aaaaanewpage aaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

I think the article should begin with following passage:

It is well known that Gödel's Incompleteness Theorem is a 2-part result, whose first half indicates no decision procedure can identify all the true statements of arithmetic whose second half specficites sufficiently strong conventional systems cannot verify their own consistency. Out title for this article may first appear to be a surprisingly long mouthful of words, but it is only manner in which it is easy to express the point our proposed boundary-case exceptions to for Gödel's Second Incompleteness Theorem are intended to be minor caveats from a theorem-proving perspective, while significant modifications if one's goals are to examine Gödel's historic millenial result from the perspective of axiomatic sytems that seek to "appreciate their own consistency" in a non-trivial semi-robust sense while "proving" their own consistency in a tender sense.

Part of the motivation of our paper can be understood when one recalls that Gödel was careful to insert into his ground-breaking paper [10], the following caveat indicating that in 1931 he considered it possible that a diluted form of Hilbert's Consistency Program could be partially successful:

- $*$: "It must be expressly noted that Proposition XI (e.g. the Second Incompleteness Theorem) 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 M or in A)."

Yourgrau has summarized, in detail, Gödel's considerations about this subject. Thus [36] indicated that "for several years" after [10]'s publication, Gödel "was cautious not to prejudge" whether some unusual formalism, different from Peano Arithmetic, might provide some type of proof of its own consistency.

It is also known that Hilbert was insistent throughout his life that axiom systems should be able to formalize some type of knowledge of their own consistency, despite Hilbert's appreciation of the significiance of Gödel's historic discovery. Thus in a context where Hilbert-Barnays developped a very important generalization of the Second Incompleteness Theorem (in the context of their three Derivabilitiy Conditions [16, 7]), Hilbert insisted throughout his life that some type of partial sucess would be ultimately obtained by his consistency program. Thus, the motto of Hilbert's Consistency Program during the 1930-1943 period was Hilbert's famous Wir mussen wissen; wir werden wissen" statement, whose English translation is given below:

- $**$: "We must know; We will know"

One notable aspect of $**$ is, of course, that it does not tell us what is the nature of the object XYZ that "we must and will know of". This was clearly because Hilbert was uncertain what XYZ was, although he insisted throughout his life that at least a revised version of his consistency program would attain at least a diluted version of success. (Indeed, history that Hilbert, who had a flare for coming up with memorable aphorisms, arranged to have statement $**$ engraved on his tombstone.) Our theme in the current article will be that the combination of self-referencing "I am consistent" axiom sentence, whose properties were described by Willard in several previous article, with what Section ??? will call "Rogue Function Definitions", can formalize a sufficiently strong XYZ to show that Hilbert's statement $**$ can be vindicated strong partial and not-full sense for one to accurately refer to axiom systems owning a surprisingly strong "appreciation" of their own consistency that still amounts to their possessing only a timid proof of their own consistency.

In a context where many articles in the scholarly world can sometimes inadvertently overstate their points, many readers may initially wonder whether the current article, whose thesis was initially developped by DW will do the same. Such skepticism is natural, and XX initially also presumed this posture before DW communicated the details of his newest results, developped in March of 2014. However in a context where the current article will score the point that there is a distinction between "proving" and "appreciatig" one's own self consistency, the latter will not occur here.

With the formalisms of google-gobble now in place, our search for the mathaphorical planet of "Altaire" from the Star Trek television series is ready to commence. These formalisms will thus serve as the Blue Angel style fuel in our coach for a Blonde Venus style planet. The above gogly-gook poetry is another way of saying that this paper will bring to completion a scientific project that I have worked on for the last 22 years.

aaaapage Even older notes below

If a proof tree has J terms and we wish to consider the blind case where all its integers are less than M and we have an atjai probing function P, then we will think of a "blind specimen as starting with a permutaion p of L integers, where $P(i) = p(i)$ and each term is mapped onto one integer. By straight forward combinatorics, the number of blind specimens in such a blind set is equal to $M^J \cdot M!$ This is a number of stunning large one, especially when one realizes that a proof of length L will have godel number $M > 2^L$.

The symbol S will denote the set of all such "blind-specimens". It will be called Starting Blind Set and thus have a stunning cardinality of $M^J \cdot M!$. This object is called a starting blind set because none of its specimens are models of an Atjai-like axiom schema because each specimen will have an object x where $F(x) = 0$.

We can also define a gigantic object into a slightly smalled obcjet called a "Reduced Blind Set". It will be the same as the Starting Blind Set except it will require each of its specimens to satisfy the following additional constraints

1) The terms t will be required to be mapped onto integers less than M in a manner that satisfies the defining rules of our functions symbols.

2) If $F^i(x)$ denotes functional object of $F(F(\ F(x)))$ (with i iterations) then none of our terms t will be mapped onto an integer x where for $i \leq J$, the object $F^i(x) = 0$

A reduced blind set R will not be quite as large S, but it will of a stunning large size greater than certainly $M^J \cdot 2^{M}!$. Each specimen of the Reduced Blind Set will violate our axioms system's requirements in a model-theoretic set, but they are still useful objects in a proof by contradiction. This is partly because each blind specimen will still be valid under our logical axioms. Moreover, if we follow a top-down path walking down our X-tableaux tree where we follow the rule that each time we reach a branch point, we will traverse the node have the largest stubset of the reduced blind set satisfying it, then each node of our path will have number of elements from R satisfying it be at least half the number of nodes satisfying all its ancestors simultaneously (except the "$\forall x \ F(x) \neq 0$" node which we are permitted to ignore.) OLD enddocument

References

ALL PAGE 5's bib references are to Willard's work
No space on page 4.

[1] Z. Adamowicz, "Herbrand Consistency and Bounded Arithmetic", Fund Mathematicae 171 (2002) pp. 279-292.
[2] Z. Adamowicz and P. Zbierski, "On Herbrand consistency in weak theories", Archive for Mathematical Logic 40 (2001) pp. 399-413.
[3] A. Bezboruah and J. Shepherdson, "Gödel's Second Incompleteness Theorem for Q", Journal of Symbolic Logic 41 (1976) 503-512.
[4] S. R. Buss, Bounded Arithmetic, (Ph D Thesis) Proof Theory Notes #3, Bibliopolis 1986.
[5] S. R. Buss and A. Ignjatovic, "Unprovability of Consistency Statements in Fragments of Bounded Arithmetic", Annals of Pure and Applied Logic 74 (1995) pp. 221-244.
[6] S. Feferman, "Arithmetization of Mathematics in a General Setting", Fund Mathematicae 49 (1960) pp. 35-92.
[7] M. Fitting, First Order Logic and Automated Theorem Proving, SpringerVerlag 1996.
[8] H. M. Friedman, "On the Consistency, Completeness and Correctness Problems", Ohio State Tech Report, 1979.
[9] H. M. Friedman, "Translatability and Relative Consistency", Ohio State Tech Report, 1979.
[10] K. Gödel, " Über formal unentscheidbare Sätze der Principia Mathematica und Verwandte Systeme I", Monatshefte für Math. Phys. 37 (1931) pp. 349-360.
[11] P. Hájek, "Mathematical Fuzzy Logic and Natural Numbers", fuzzy logic", Fundamenta Mathematicae 81 (2007) pp. pp 155-163 (with related results also appearing in reference [12].
[12] P. Hájek, "Towards metamathematics of weak arithmetics over fuzzy logic", Logic Journal of the IPL 19 (2011) pp. 467-475.
[13] P. Hájek and P. Pudlák, Metamathematics of First Order Arithmetic, Springer Verlag 1991.
[14] D. Hilbert, Uber das Unendliche, Springer 1926. (English Translation (1967) by J V Heijenoort in From Frege to Gödel: A Source Book in Math Logic, Harvard Press, page 376)
[15] D. Hilbert and P. Bernays, Grundlagen der Mathematik, Springer 1939.
[16] R. Jeroslow, "Consistency Statements in Formal Mathematics", Fundamenta Mathematicae 72(1971) 17-40.
[17] S. C. Kleene, "On the Notation of Ordinal Numbers", Journal of Symbolic Logic 3 (1938), 150-156.
[18] J. Krajícek, Bounded Propositional Logic and Complexity Theory, Cambridge University Press, 1995.
[19] L. A. Kołodziejczyk, "On the Herbrand Notion of Consistency for Finitely Axiomatizable Fragments of Bound Arithmetic", Journal of Symbolic Logic 71 (2006) pp. 624-638.
[20] G. Kreisel and G. Takeuti, "Formally Self-Referential Propositions for Cut-Free Classical Analysis", Dissertationes Mathematicae 118 (1974) pp. 1-55
[21] Richard Lipton's May 21, 2009 Blogging Column REMAINDER IN COMMENT LINE TO SAVE SPACE ...
[22] M. H. Löb, A Solution to a Problem by Leon Henkin, Journal of Symbolic Logic 20(1955) 115-118
[23] E. Nelson, Predicative Arithmetic, Princeton Math Notes, 1986.
[24] R. Parikh, "Existence and Feasibility in Arithmetic", Journal of Symbolic Logic 36 (1971) 494-508
[25] J. B. Paris and C. Dimitracopoulos, "A Note on the Undefinability of Cuts", Journal of Symbolic Logic 48 (1983) pp. 564-569.
[26] C. Parsons, On n-Quantifier Elimination, Journal of Symbolic Logic 37 (1972), pp. 466-482.
[27] P. Pudlák, "Cuts, Consistency Statements&Interpretations", Journal of Symbolic Logic 50(1985)423-442
[28] P. Pudlák, "On the Lengths of Proofs of Consistency", in Collegium Logicum: 1996 Annals of the Kurt Gödel Society (Volume 2), Springer-Wien-NewYork, pp 65-86.
[29] H. Rogers Jr., Recursive Functions and Effective Compatibility, McGrawHill 1967.
[30] R. M. Solovay, "Provability Interpretations in Modal Logic", Israel Journal of Mathematics 25 (1976) pp. 279-304.
[31] R. M. Solovay, "Injecting Inconsistencies into Models of PA", Annals Pure and Applied Logic" 44(1989) 102-132
[32] R. M. Solovay, Private telephone communications during April of 1994 REMAINDER IN COMMENT LINE TO SAVE SPACE ...
[33] V. Švejdar, "An Interpretation of Robinson Arithmetic in its Grzegorczjk's Weaker Variant" Fundamenta Informaticae 81 (2007) pp. 347-354.
[34] A. Visser, "The Unprovability of Small Inconsistency", Archive for Mathematical Logic 32 (1993) pp. 275-298.
[35] A. J. Wilkie and J. B. Paris, "On the Scheme of Induction for Bounded Arithmetic", Annals of Pure and Applied Logic (35) 1987, 261-302
[36] P. Yourgrau, A World Without Time: The Forgotten Legacy of Gödel and Einstein, Basic Books (2005). (Our quotations from [36] come from its page 58.)
[37] D. Willard, "Self-Verifying Axiom Systems", Proceedings of the Third Kurt Gödel Colloquium (1993), Springer-Verlag LNCS#713, pp. 325-336.
[38] D. Willard, "Self-Reflection Principles and NP-Hardness", DIMACS Series in Discrete Mathematics and Theoretical Computer Science # 39 (1996), pp. 297-320.
[39] —, "The Semantic Tableaux Version of the Second Incompleteness Theorem Extends to Robinson's Arithmetic Q", Proceedings of the Tableaux 2000 Conference, SpringerVerlag LNAI#1847, 415-430

Dan Willard signed before me in Glenmont, NY on 5/10/14.

BERTON
e of New York
6172156
ny County
Aug. 6, 2015