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 inevetibly 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 [5, 33].

Their results establish that logical axiom systems, employing Hilbert-style deductive methods, obey more powerful generalisations of the Second *Incompleteness* Theorem than the traditional classic generalisations 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 notarised until May10 because my Wollic paper was accepted on May 2 (and rewrting it required my attention).

NEW *Section's Title Background Notation*

Throughout this article, α will denote a recursively enumerable set of "proper axioms", and d will denote a deduction method, used for deriving theorems from α. 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 α'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 α, shall greatly simplify several aspects of our proof. We therefore emphasise that by the axiom system " α ", 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 any of Enderton, Mendelson or Papadimitriou [?, ?, ?]. 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 [?] as a framework to define L_d. (This framework was also employed in Papadimitriou's textbook [?]).

2. Mendelson's framework [?], 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 [?]) 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 mostly 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 disadvantadge, 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 muster 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 α denote an axiom system, and d denote a deduction method. The ordered pair (α, d) will be called Self Justifying when:

i one of α 's theorems (or at least one of its axioms) 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.

Example 0.4 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(α, d)" (looking at itself).

Kleene [17] discussed how to encode approximate analogs of this "SelfRef(α, d)" statement. Each of Kleene, Rogers and Jeroslow [17, 29, 16] noted α^d may, however, be inconsistent (despite SelfRef(α, d)'s assertion), thus causing it to violate Part-ii of self-justification's definition. This is because if the ordered pair (α, d) is too strong then a classic Gödel-style diagonalisation argument can be applied to the axiom system $\alpha^d = \alpha +$ SelfRef(α, d), where the added presence of the statement SelfRef(α, d) causes this extended version of α to ironically become automatically inconsistent. Thus, the machinery of the sentence "SelfRef(α, d)", while relatively easy to encode via an application of the Fixed Point Theorem, is ironically *most often useless!*

We are now ready to explain the intution 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 italicised 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 *reason 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 $x * y = z$. An axiom system α will be said to recognise 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 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. Moreover, α 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 formalise 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, summarised by Theorem 3.1, 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 α denote any consistent Type-S system that can verify addition and multiplication satisfy their usual associative, commutative, distributive and identity axioms. Then α 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+ *year effort to prove $P \neq NP$* would have taken more then 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 receive tenure from Berkeley's Math Department before he published a single paper.* Also, many logicians [5, 13, 18, 28, 37, 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 [40] 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 [8] 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 taboo "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 of 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

Dan E Willard

Jodi Sherman Schultz
No. 01SH6181014
Notary Public, State of New York
Qualified in Albany County
Commission Expires 01/22/2016

It turns out that there are many different variants of Type-NS self-verifying 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 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 discourse, 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 Semantics" 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) technically 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:

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

We will assume the "name" of the built-in constant symbol "C_j" is encoded using $O[\,\mathrm{Log}(j+2)\,]$ 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 "$C_5 - C_3 - C_2$".

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 generalization of the Second Incompleteness Theorem will apply to settings where one replaces Equation (4)'s additive naming convention with (5)'s "Multiplicative Naming Convention" :

$$\mathrm{Mult}\,(C_j^*\,,\,C_j^*\,,\,C_{j+1}^*\,) \tag{5}$$

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 $\mathrm{Pred}(x) = \mathrm{Min}(x-1, 0)$, as well as contains a constant symbol C_1 for representing the value "1". Also, let $\mathrm{Pred}^j(x)$ 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 $\mathrm{Pred}^{k-1}(k) = C_1$). 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).

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

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 hybridization 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 conceptualize, 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.
ddd2.3

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 generalisations 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:*