Notre Dame Journal of Formal Logic
Volume 26, Number 4, October 1985

1-Consistency and the Diamond

GEORGE BOOLOS

1 Introduction It is well known that the set of (Gödel numbers of) sentences of arithmetic that are consistent with classical first-order arithmetic with induction (Peano Arithmetic (*PA*)) is Π-1 complete. Solovay showed in [5] that the propositional modal logic characterizing consistency in *PA* is the system of modal logic known variously as *G*, *GL*, and *L*: the formulas of modal logic that are theorems of *G* are precisely those that are provable in *PA* under all substitutions of sentences of arithmetic for atoms p_0, p_1, $p_2 \ldots$ of modal logic, the diamond $\Diamond$ and box $\Box$ of modal logic being respectively interpreted as the consistency predicate $Con(x)$ and the provability predicate $Bew(x)$ of arithmetic.[1] In [3], building upon Solovay's work, I showed that the set of sentences that are ω-consistent with *PA* is Π-3 complete and that the system *G* is also the modal logic characterizing ω-consistency. Thus despite the greater complexity of its definition, there is a natural and easily definable class of properties in respect of which ω-consistency does not differ from (simple) consistency.

A theory *T* in the language of arithmetic is said to be *1-inconsistent* if for some *primitive recursive* formula Rx, T implies $\exists x{-}Rx$ and also implies $R\mathbf{n}$, for every natural number n; *T* is 1-consistent if it is not 1-inconsistent. The definition of ω-consistency thus differs from that of 1-consistency only in lacking the qualifier "primitive recursive". Obviously, every ω-consistent theory is 1-consistent and every 1-consistent theory is consistent; neither converse holds.

1-consistency was first defined by Kreisel. Some interesting facts about it are: (i) A modification of the finite version of Ramsey's theorem, due to Paris and Harrington, turns out to be equivalent in *PA* to the assertion of 1-consistency, as do a number of other "mathematically interesting, non-self-referential" undecidable sentences devised by other authors. (ii) In his proof of the incompleteness theorems, Gödel constructed a sentence which he showed to be undecidable in the system under consideration on the assumption that the system is ω-consistent. As Kreisel observed, however, this assumption is unnecessarily strong; the assumption that the system is 1-consistent suffices to show that the sentence Gödel constructed is undecidable. (Rosser showed that a certain

Received February 1, 1984; revised July 23, 1984

other sentence could be shown undecidable on the assumption of simple consistency.) (iii) The assumption of 1-consistency decides the truth or falsity of every sentence built up from $0 = 1$ by means of truth-functional connectives and the formula $Bew(x)$, among which are such sentences as $-Bew(\ulcorner 0 = 1 \urcorner)$, which is the consistency assertion, and $(Bew(\ulcorner -Bew(\ulcorner 0 = 1 \urcorner) \urcorner) \rightarrow Bew(\ulcorner 0 = 1 \urcorner))$.

A *sentence S* is said to be 1-consistent with a theory T if the theory whose axioms are S and the axioms of T are 1-consistent. Henceforth $T = PA$ and reference to PA will always be tacitly understood. The set of (Gödel numbers of) sentences that are {ω-/1-} consistent is at worst Π-3: S is {ω-/1-} consistent if *for all* {formulas/primitive recursive formulas} Rx, either there is no proof of $(S \rightarrow \exists x - Rx)$ or *for some* natural number n, *for every* proof p, $(S \rightarrow R\mathbf{n})$ is not the last line of p. But although the classification of the set of ω-consistent sentences cannot be improved since this set is Π-3 *complete*, that of the set of 1-consistent sentences can: we shall show that this set is Π-2, and indeed Π-2 complete.

Furthermore, we shall show that the system G is also the modal logic characterizing the notion 1-consistency. Thus consistency, 1-consistency, and ω-consistency are Π-1, Π-2, and Π-3 complete, respectively, but all have the same propositional modal logic.

2 **Π-*2 completeness*** A *formula F* is called Π-1 if $F = \forall x Rx$, for some primitive recursive formula Rx (which may contain free variables other than x). F is Σ-2 if $F = \exists x \forall y Rxy$, for some primitive recursive Rxy.

The following well-known characterization of the notion of a 1-consistent sentence is the key step in the proof of the Π-2-ness of 1-consistency.

Theorem 1 *A sentence S is 1-consistent iff S is consistent with every true* Π-1 *sentence.*

Proof: Suppose that Rx is a primitive recursive formula, $\forall x Rx$ is true, and $(S \,\&\, \forall x Rx)$ is inconsistent. Then for every n, $R\mathbf{n}$ is true, $R\mathbf{n}$ is provable (by the Σ-1 completeness of PA), and thus S implies $R\mathbf{n}$. But S also implies $\exists x - Rx$ and S is therefore 1-inconsistent.

Conversely, suppose S is 1-inconsistent. Then for some primitive recursive Rx, S implies $\exists x - Rx$ and implies $R\mathbf{n}$ for every n. If $R\mathbf{n}$ is true for every n, then S is inconsistent with the true Π-1 sentence $\forall x Rx$; if for some n, $R\mathbf{n}$ is not true, then $-R\mathbf{n}$ is true, hence provable by Σ-1 completeness, and S is outright inconsistent.

Theorem 2 *The set of (Gödel numbers of) 1-consistent sentences is* Π-2 *complete.*

Proof: Since the set of true Π-1 sentences and the set of consistent sentences are both Π-1, the set X of sentences S such that for every U (if U is a true Π-1 sentence, then $(U \,\&\, S)$ is consistent) is visibly Π-2. But by Theorem 1, X is the set of 1-consistent sentences.

Let Y be a Π-2 set. We must show how to reduce Y to X. Since Y is Π-2, there is a primitive recursive formula $Hxyz$ such that for any natural number

m, m is in Y iff $\forall y \exists z H \mathbf{m} yz$ is true in the standard model. We'll show that m is in Y iff $\forall y \exists z H \mathbf{m} yz$ is 1-consistent.

Suppose m is not in Y. Then for some n, $\forall z\text{-}H\mathbf{mn}z$ is a true Π-1 sentence with which $\forall y \exists z H \mathbf{m} yz$ is inconsistent. By Theorem 1, $\forall y \exists z H \mathbf{m} yz$ is 1-inconsistent. Conversely, if m is in Y, then $\forall y \exists z H \mathbf{m} yz$ is true and therefore certainly 1-consistent.

3 The characterization result We shall now show that G is the modal logic of 1-consistency. Instead of working directly with the notion of 1-consistency we shall work with the more convenient dual notion of 1-provability (provability from some true Π-1 sentence). To prove the analogue of Solovay's theorem for 1-consistency, we must first establish the analogues for 1-consistency of the usual derivability conditions. Some preliminaries:

Definition A natural number m is (the Gödel number of) a *1-proof* of a sentence S if m is (the Gödel number of) a proof of a conditional whose consequent is S and whose antecedent is some true Π-1 sentence.

Definition A sentence S is *1-provable* (or a *1-theorem*) if some m is a 1-proof of S.

Lemma 1 *S is 1-provable iff -S is 1-inconsistent.*

Proof: Theorem 1 and De Morgan.

We write "$\vdash S$" as an abbreviation of "S is provable (in PA)". A dictionary of terms and formulas:

$Pf(y)$: the usual primitive recursive formula for the set of Gödel numbers of proofs.

$LL(y)$: a primitive recursive term for λj (the Gödel number of the last line of the proof whose Gödel number is j if j is the Gödel number of a proof, else 0).

$Pf(y, x)$: the formula $(Pf(y) \mathbin{\&} x = LL(y))$.

$Bew(x)$: the formula $\exists y Pf(y, x)$.

$\ulcorner F \urcorner$: the numeral for the Gödel number of F.

$Sub(x, y, z)$: a primitive recursive term for λijk (the Gödel number of the result of substituting the expression with Gödel number i for the jth variable in the expression with Gödel number k).

x: the first variable.

$Num(x)$: a primitive recursive term for λi (the Gödel number of the numeral for i).

$Tr(z)$: the usual Π-1 satisfaction formula for Π-1 formulas, with the property that for any Π-1 formula F, the biconditional $(F \leftrightarrow Tr(Sub(Num(x), 1, \ulcorner F \urcorner)))$ is provable.[2]

$Ante(z)$: a primitive recursive term for λk (the Gödel number of the antecedent of the formula with Gödel number k if k is the Gödel number of a conditional, else 0).

$Cons(z)$: a primitive recursive term for λk (the Gödel number of the consequent of the formula with Gödel number k if k is the Gödel number of a conditional, else 0).

1-$Pf(y, x)$: the formula $(Pf(y)\ \&\ Tr(Ante(LL(y)))\ \&\ x = Cons(LL(y)))$. (Notice that 1-$Pf(y, x)$ is (equivalent to) a Π-1 formula.)

1-$Bew(x)$: the formula $\exists y$1-$Pf(y, x)$. (1-$Bew(x)$ is a Σ-2 formula.)

We can now demonstrate analogues of the derivability conditions for 1-$Bew(x)$.

Lemma 2 *If* $\vdash S$, *then* S *is* 1-*provable.*

Proof: If $\vdash S$ then $\vdash(\forall x x = x \rightarrow S)$.

Lemma 3 $\vdash Bew(\ulcorner S \urcorner) \rightarrow$ 1-$(Bew(\ulcorner S \urcorner)$.

Proof: Formalize Lemma 2.

Lemma 4 *Suppose* $\vdash S$. *Then* $\vdash$1-$Bew(\ulcorner S \urcorner)$.

Proof: If $\vdash S$, then by one of the derivability conditions for $Bew(x)$, $\vdash Bew(\ulcorner S \urcorner)$, whence by Lemma 3, $\vdash$1-$Bew(\ulcorner S \urcorner)$.

Lemma 5 *If* S *and* $(S \rightarrow S')$ *are* 1-*provable, so is* S'.

Proof: If S and $(S \rightarrow S')$ are implied by true Π-1 sentences P and Q, then S' is implied by $(P\ \&\ Q)$, which is (equivalent to) a true Π-1 sentence.

Lemma 6 $\vdash$1-$Bew(\ulcorner S \urcorner)$ & 1-$Bew(\ulcorner (S \rightarrow S') \urcorner) \rightarrow$ 1-$Bew(\ulcorner S' \urcorner)$.

Proof: Formalize Lemma 5.

Lemma 7 *Let* S *be a* Σ-2 *sentence. Then* $\vdash S \rightarrow$ 1-$Bew(\ulcorner S \urcorner)$.

Proof: Let $S = \exists x \forall y Rxy$, Rxy primitive recursive. For any natural number n, let $r(n)$ be the result of substituting the numeral for n for the variable x in $\forall y Rxy$, i.e., $r(n) = \forall y R\mathbf{n}y$. Now formalize the following argument, using the terms and formulas in the dictionary: if S holds, then for some n, $r(n)$ is a true Π-1 sentence. For every n, the conditional with antecedent $r(n)$ and consequent S is provable. Therefore if S holds, S is 1-provable.

Lemma 8 $\vdash$1-$Bew(\ulcorner S \urcorner) \rightarrow$ 1-$Bew(\ulcorner$1-$Bew(\ulcorner S \urcorner)\urcorner)$.

Proof: By Lemma 7 and the fact that 1-$Bew(x)$ is Σ-2.

We turn now to the connection with the systems G and G^* of modal logic. The axioms of the system G are all tautologies, all sentences $(\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B))$, and all sentences $(\Box(\Box A \rightarrow A) \rightarrow \Box A)$; the rules of inference of G are necessitation (if A is derivable, then so is $\Box A$) and modus ponens. The axioms of the system G^* are the theorems of G and all sentences $(\Box A \rightarrow A)$; the sole rule of inference of G^* is modus ponens.

Let "f" be a variable ranging over functions from the atoms of modal logic to sentences of arithmetic. For each sentence A of modal logic, define Af by: $Af = f(A)$ if A is an atom; f commutes with propositional connectives; and $(\Box A)f = Bew(\ulcorner Af \urcorner)$. (Thus $(\Diamond A)f$ is equivalent to the sentence of arithmetic formalizing the assertion that Af is consistent.) Similarly, define A^1f by: $A^1f = f(A)$ if A is an atom; $(-A)^1f = -(A^1f)$; $(A\ \&\ B)^1f = (A^1f\ \&\ B^1f)$; and similarly for the other propositional connectives; and $(\Box A)^1 f = 1\text{-}Bew(\ulcorner A^1 f \urcorner)$. $((\Diamond A)^1 f$ is then equivalent to the sentence of arithmetic formalizing the assertion that $A^1 f$ is 1-consistent.)

The two completeness theorems of Solovay for G and G^* are that a modal formula A is a theorem of G iff Af is provable for all f, and that a modal formula is a theorem of G^* iff Af is true for all f. Artyomov, Leivant, Montagna, and the author have proved an extension of Solovay's completeness theorem for G: there is a *fixed* f such that for all modal formulas A, if A is not a theorem of G, then Af is not provable. Of course, since either $p_0 f$ or $-p_0 f$ is true, there can be no such "uniform" analogue for the system G^*. We shall now show how to establish analogues of all these results for the notion of 1-consistency.

Our main theorem is the following:

Theorem 3 *For every A, f, if A is a theorem of G, then $A^1 f$ is provable for all f; for some f, for every A, if A is not a theorem of G, then $A^1 f$ is not provable; and for every A, A is a theorem of G^* iff $A^1 f$ is true for all f.*

Proof: Together with the diagonal lemma, Lemmas 4, 6, and 8 ensure the provability in PA of the analogue for 1-provability of Löb's theorem. It is then clear that for every f, $A^1 f$ is provable if A is a theorem of G, and (since a 1-provable statement is true) that $A^1 f$ is true if A is a theorem of G^*.

We now show how to amend the proof of the main theorem given in [2], viz. that for some f, for all A, if A is not a theorem of G, then Af is not provable, so that it becomes a proof of the analogous result for 1-consistency. We indicate the changes that must be made in pp. 192–195 of that paper in order to establish the analogous result. First of all, replace "ϕ" everywhere by "f". Then inset "1-"s before "f", "Pf", "Bew", "proof", "theorem" at the appropriate places. Replace mention of the Hilbert-Bernays derivability conditions by reference to Lemmas 4, 6, and 8 above; mention of provable Σ-1 completeness, by reference to Lemma 7. The conclusion of the proof should read, "$-1\text{-}Bew(\ulcorner A^1 f \urcorner)$ is true, and thus $A^1 f$ is not a 1-theorem of PA, and hence not a theorem of PA".

The most noteworthy change, however, is in the definition on p. 193 of $\theta(x_1, x_2)$, which is no longer a primitive recursive term, but, in view of the Π-1-ness of "1-proof", a Σ-2 term for the function whose value at m, r is j if r is the Gödel number of a 1-proof of the formula mentioned in the original definition of θ, and is 0 otherwise, i.e., there is a Σ-2 formula $s(x_1, x_2, z)$ that is satisfied by the graph of this function and is such that $\forall x_1 \forall x_2 \exists! z s(x_1, x_2, z)$ is provable. The new $G(a, b)$ is thus no longer a Σ-1 formula, but rather a Σ-2 formula, as can be seen by rewriting $B(y, a, b)$ as:

$$\begin{aligned} &\exists x(Lh(s) = a + 1 \;\&\; (s)_0 = 0 \;\&\; (s)_a = b \;\&\\ &\quad \forall x < a\{[\forall z(s(y, x, z) \rightarrow \rho((s)_x, z)) \rightarrow \exists z s(y, x, (s)_{x+1})] \;\&\\ &\qquad [-\exists z(s(y, x, z) \;\&\; \rho((s)_x, z)) \rightarrow (s)_{x+1} = (s)_x]\}) \;. \end{aligned}$$

These changes made, the proofs of analogues of (A)–(E), the lemma, and the main theorem, proceed exactly as in [2].

As for G^*, the result that if A is not a theorem of G^*, then for some f, A^1f is not true, may be obtained by making a similar, routine, modification of the proof of Solovay's theorem for G^* given in Chapter 12 of [1].

4 Final remarks

1. Although it is not in general the case that A is a theorem of G iff Af is 1-provable for all f (let $A = \Diamond T$), our proof and Lemma 2 above show that A is a theorem of G iff A^1f is 1-provable for all f.

2. We call a sentence S of arithmetic *extremely undecidable* if for all A containing no atom other than p_0, if Ag is not provable for some g, then neither is Af for any f such that $f(p_0) = S$. No Σ-1 or Π-1 sentence can be extremely undecidable. In [2], we showed the existence of infinitely many Δ-2 extremely undecidable sentences. In fact, if we define $F(y) = \exists x(S(x) \mathbin{\&} \Pi(x, y))$, where $S(x)$ and $\Pi(x, y)$ are as in that paper, then $F(y)$ is a Δ-2 predicate such that if $f(p_i) = F(\mathbf{i})$ for all i, then A is a theorem of G iff Af is provable; thus the numerical instances of $F(y)$ are Δ-2 extremely undecidable sentences. If we define "extremely 1-undecidable" analogously, then no Δ-2 or Π-2 sentence can be extremely 1-undecidable, but we may use the devices of [2] to conclude that there is a Δ-3 predicate $F'(y)$ such that if $f(p_i) = F'(\mathbf{i})$ for all i, then A is a theorem of G iff A^1f is provable.

3. An entirely parallel treatment can be given for ω-consistency. By introducing the Π-2 notion of an *ω-proof* of S, i.e., a proof of a sentence $(\forall xFx \rightarrow S)$ such that for all n, $F\mathbf{n}$ is provable, noticing that $-S$ is omega-inconsistent iff there is an ω-proof of S, and using a Π-2 formula to formalize *ω-proof*, one can prove an analogue of Theorem 3 for omega-consistency. (For further details see [3].) Moreover, there is a Δ-4 predicate with properties analogous to those of $F(y)$ and $F'(y)$ above.

NOTES

1. [1] provides an account of the relation between the system G and the concepts of provability and consistency in *PA*. Notation and terminology not defined in this paper are explained in that work.

2. Cf. p. 843 of [4].

REFERENCES

[1] Boolos, G., *The Unprovability of Consistency*, Cambridge University Press, Cambridge, 1979.

[2] Boolos, G., "Extremely undecidable sentences," *The Journal of Symbolic Logic*, vol. 47 (1982), pp. 191–196.

[3] Boolos, G., "ω-consistency and the diamond," *Studia Logica*, vol. 39 (1980), pp. 237–243.

[4] Smoryński, C., "The incompleteness theorems," pp. 821–865 in *Handbook of Mathematical Logic*, ed., Jon Barwise, North-Holland, Amsterdam, 1977.

[5] Solovay, Robert, "Provability interpretations of modal logic," *Israel Journal of Mathematics*, vol. 25 (1976), pp. 287–304.

Department of Linguistics and Philosophy
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139