Logic and **L**ogical **P**hilosophy
Volume 6 (1998), 63–76

Fabrice Pataut

INCOMPLETENESS, CONSTRUCTIVISM AND TRUTH

Abstract. Although Gödel proved the first incompleteness theorem by intuitionistically respectable means, Gödel's formula, true although undecidable, seems to offer a counter-example to the general constructivist or anti-realist claim that truth may not transcend recognizability in principle. It is argued here that our understanding of the formula consists in a knowledge of its truth-conditions, that it is true in a minimal sense (in virtue of a *reductio ad absurdum*) and, finally, that it is recognized as such given the consistency and ω-consistency of P. The philosophical lesson to be drawn from Gödel's proof is that our capacities for justification in favour of minimal truth exceed what is strictly speaking formally provable in P by means of an algorithm.

Received June 20, 1998

Does Gödel's first incompleteness theorem have consequences for the question whether intuitionistic semantics should be preferred to classical semantics and, given that the acceptance of semantic principles entails the acceptance of corresponding logical laws, for the question whether intuitionistic logic should be preferred to classical logic?

Although some non-constructive proofs of the theorem have been proposed after the publication of Gödel's 1931 result, e.g. by Boolos (1989), it may seem obvious that Gödel's original proof cannot have any bearing on the issue of the choice of logic, for it remains conspicuously neutral between the classical and the intuitionistic standpoints. Gödel thought it worthwhile to remind his readers that his proof was indeed constructive. He pointed to the fact that the first incompleteness result had been obtained "in an intuitionistically unobjectionable manner" (Gödel [1931: 189], 1986a: 177) and offered as a warrant for his claim that "all existential statements [*Existentialbehauptungen*] occurring in the proof [were] based upon Theorem V [i.e. the theorem immediately preceding the first incompleteness theorem], which, as is easily seen, is unobjectionable from the intuitionistic point of view" (Gödel *loc. cit.*: note 45a).

In Kleene's terminology, Theorem V states that every primitive recursive relation is numeralwise expressible in P, where P is the system obtained from Whitehead and Russell's *Principia Mathematica*, without the ramification of the types, taking the natural numbers as the lowest type and adding their usual Peano axioms (Kleene 1986: 132). When expressed formally, without reference to any particular interpretation of the formulas of P, and in Gödel's own terminology, which favours the indirect talk of 'Gödel' numbers and concepts applying to those numbers rather than a direct talk of the formal objects, Theorem V claims that:

GÖDEL'S ([1931], 1986a) THEOREM V
For every recursive relation $R(x_1, \ldots, x_n)$ there exists an n-place RELATION SIGN r (with the FREE VARIABLES u_1, u_2, …, u_n) such that for all n-tuples of numbers $(x_1, \ldots, x_n)$ we have

$$\begin{aligned} R(x_1, \ldots, x_n) &\rightarrow \mathrm{Bew}[\mathrm{Sb}(r^{u_1 \ldots u_n}_{Z(x_1) \ldots Z(x_n)})], \\ \overline{R}(x_1, \ldots, x_n) &\rightarrow \mathrm{Bew}[\mathrm{Neg}(\mathrm{Sb}(r^{u_1 \ldots u_n}_{Z(x_1) \ldots Z(x_n)}))]. \end{aligned}$$

Gödel gives an outline of the proof and notes, on this occasion, that Theorem V is itself "of course, [...] a consequence of the fact that in the case of a recursive relation R it can, for every n-tuple of numbers, be decided *on the basis of the axioms of the system P* whether the relation R obtains or not" (Gödel *op. cit.*: [186n39], 171n39). This, it must be noted, can also be decided by means of procedures which remain unobjectionable from the intuitionistic standpoint.

One may object that, as far as a choice in favour of a given semantics is concerned, either classical or otherwise, it hardly matters whether or not Gödel's proof is intuitionistically safe *logically* speaking, for, although the acceptance of semantic principles normally entails the acceptance of corresponding logical laws, the converse does not hold.[1] If this conception of the relation between semantic principles and logical laws is correct, Gödel could very well have used, say, the law of excluded middle in carrying out his proof without thereby committing himself to the principle of bivalence (every statement is either true or false) ; or he could have used the law of double negation elimination without thereby committing himself to the principle of stability (every statement which is not false is true). If the remark applies to fundamental logical laws and fundamental semantic principles quite generally, and not only to excluded middle and double negation elimination, then the non-constructive proofs of the theorem should not imply any semantic claim which a constructivist or intuitionist would have to reject. The problem is that they do indeed imply such claims. Boolos' proof, in particular, establishes the existence of an undecidable statement of arithmetic, just like Gödel's ; but, unlike Gödel's, it does not provide an effective procedure for producing it. Let a *correct* algorithm M be an algorithm which may not list a false statement of arithmetic. A truth omitted by M is just a true sentence of arithmetic not listed by M. Boolos' proof establishes the existence of such a true statement, but the statement is recognized to be true only classically and not constructively.

The problem, now, is whether the *rejection* of a given logical law entails the rejection of the corresponding semantic principle. Of course, *if* the acceptance of, say, bivalence across the board entails the endorsement of excluded middle, then, by contraposition, the rejection of excluded middle entails the rejection of bivalence. If acceptance goes one way, from the semantic to the logical, then rejection must go the other way, from the logical to the semantic. Maybe we would like to treat logical laws and semantic

[1] The point has been made, e.g., by Michael Dummett in Dummett (1978: xix).

principles on a par in the following sense: just as the acceptance of a given semantic principle entails the acceptance of a corresponding logical law, the rejection of the same semantic principle should entail the rejection of the same corresponding logical law. But, clearly enough, the order of endorsement and the order of rejection may not be the same.

There is something deeply unsatisfying in this situation, for if we adopt the point of view according to which, very roughly speaking, one must argue from semantics to logic and absolutely not the other way around, we may find ourselves in a situation where, even if we refrain from using classical reasoning in our proof of Gödel's first incompleteness theorem, we might nevertheless end up endorsing the classical semantic principles which an intuitionnist should reject. We would certainly be in a quite uncomfortable situation *philosophically* speaking if Gödel, or anyone proposing an intuitionistically acceptable proof of the first incompleteness theorem, nevertheless contravened the fundamental intuitionist, constructivist, or anti-realist *semantic* claim according to which truth must be, at least in principle, acknowledgeable, recognizable in some way or other as obtaining whenever it does indeed obtain.

Gödel's true although undecidable formula, the existence of which is proved constructively by the first incompleteness theorem, does offer, at least at first sight, a counter-example to the constructivist or anti-realist semantic claim according to which truth may not transcend recognition, or at least recognition in principle, for the proof of the theorem establishes the existence of a formula which does have both properties, that of truth and that of undecidability. Gödel's diagonal argument does indeed provide a true statement which is nevertheless omitted by the relevant algorithm.

Two aspects of the situation passed on to us by Gödel's proof and result somewhat complicate the matter. First, there are followers of Wittgenstein's *Remarks on the Foundations of Mathematics*, like Shanker (1989), who think that it is simply incoherent and even nonsensical to claim that a statement or formula is both true and undecidable (or even that it may be). Shanker points out that, when a complete manifestation of our recognition of the truth of the Gödel formula should be made possible, the major defect of the semantic formulation of the theorem is that it turns the connection between a mathematical statement and its proof into a purely external matter (Shanker *op. cit.*: 221 ff.). This strongly suggests that the first incompleteness theorem should be formulated in a purely syntactical manner, without reference or commitment to truth, as stating, strictly speaking, that every formal system S, when elementary number theory is taken as its domain, if it is consistent, contains a formula $\mathscr{A}$ which expresses a proposition $\boldsymbol{A}$

of elementary number theory, such that neither $\mathscr{A}$ nor its negation $\neg\mathscr{A}$, expressing $\neg\boldsymbol{A}$, is provable in any of these systems.

Secondly, there is Gödel's peculiar conception — peculiar, of course, in view of the fact that classical logic and intuitionistic logic come into deep conflict over the meaning of the logical constants — that "[...] intuitionistic logic, as far as the calculus of propositions and of quantification is concerned, turns out to be rather a renaming and reinterpretation than a radical change of classical logic" (Gödel [1941: 3], 1995: 190). In particular, there is the surprising conception according to which the law of excluded middle is intuitionistically acceptable.

In classical propositional logic, '$\sim$' being taken as the classical negation sign, the formula '$p \vee \sim p$' is a tautology. According to Gödel (*op. cit.*: [2], 190), it is sufficient to define a notion of disjunction such that,'$\neg$' being taken as the intuitionistic negation sign, '$p \vee \neg p$' is also a tautology. Gödel proposes to define '$p \vee q$' as '$\neg(\neg p \ \& \ \neg q)$' — the equivalence of the two schema being known as the fourth de Morgan's law — , '$p \vee \neg p$' becoming thus '$\neg(\neg p \ \& \ \neg\neg p)$', the law of excluded middle being thereby nothing more than a special case of the law of non-contradiction, which is, of course, intuitionistically valid.

Gödel ([1933], 1986b), building on results by Glivenko (1929), showed that the classical propositional calculus is a subsystem of the intuitionistic propositional calculus and that every valid classical formula also holds in Heyting's propositional calculus *provided* that we translate the following classical notions:

Classical logical constants

1	2	3	4
$\sim p$	$p \rightarrow q$	$p \vee q$	$p \cdot q$

into the following intuitionistic notions:

Intuitionistic logical constants

1	2	3	4
$\neg p$	$\neg(p \ \& \ \neg q)$	$\neg(\neg p \ \& \ \neg q)$	$p \ \& \ q$

Of course, such a translation overlooks the fact that some intuitionists, whom we may call 'exclusivists' with respect to the logic they advocate, reject classical logic precisely because they think that classical deduction rules allow *illegitimate* inferences. What these intuitionists will reject, while still holding to the law of non-contradiction, is the idea that either p or its negation could be true whether or not we could ever be able to detect it. They will object to the idea that one may infer the negation of p from the absurdity of the supposition that p could be true independently of a proof. Typically, they will refuse to assert the negation of p unless we obtain a *reductio ad absurdum* of the supposition that we could obtain a *proof* of p. So there is, at a quite fundamental level, a deep disagreement concerning the logical form which a *reductio* should take. A classical *reductio* is not logically equivalent to an intuitionistic *reductio*. Each yields a particular form of negation which cannot be reduced to the other. This suggests that, contrary to what the first column of the translation manual suggests, '$\sim$' and '$\neg$' may not have the same meaning. Any intuitionist taking an exclusivist stance with respect to the logic he advocates will reject both the classical law of excluded middle, '$p \vee \sim p$', and the proposed translation. His rationale for the rejection will be that the meaning imposed on the disjunction and negation signs by the translation manual is such that a claim to the effect that '$p \vee \neg p$' does *not*, despite the fact that the schema contains an occurrence of the intuitionistic negation sign, amount to the claim that we have either proved p or its negation, or that we could at least be in a position to do so.

In this connection, one may also look at Brouwer's claim that one could very well endorse the logical law of excluded middle without thereby being committed to the semantic principle of bivalence (Brouwer [1948] 1983). Brouwer remarks that intuitionistically speaking, 'p is true' just means 'p is proved' or 'p may be proved'. Its intuitionistic contradictory is not 'p is false' but 'p is not true', by which an intuitionist must mean 'p is not proved to be true' and not 'p is proved to be not true'. But this shows that the law of excluded middle does not entail bivalence *only* provided that we have replaced classical truth by provability in principle or warranted assertability. Such a replacement certainly involves a conception of negation, disjunction and truth which is weaker than the conception which a classicist with respect to semantic principles and logical laws would care to defend. It yields a principle according to which either we have a proof that p is true, or we have a proof that p has been reduced to absurdity, and *not* a principle according to which either p is true or p is false.

The point I wish to make here is semantic in nature. It does not rest on any particular conception of the relation between semantic principles and logical laws and it does not concern bivalence, at least not directly. My purpose is to show that, although it is legitimate to claim that Gödel's formula is true, the first incompleteness result leaves the debate over the legitimacy of the modal schema '$\Diamond$(p is true & p is undecidable)' and its various instances unscathed. One may not, on the basis of Gödel's result, argue that arithmetical truth may transcend all possible recognition.

To begin with, is it legitimate to claim that Gödel's undecidable formula is true *simpliciter*? When giving an informal sketch of the proof in the first section of his 1931 article, Gödel ([1931: 175], 1986a: 149) says that if the proposition $[R(q); q]$ were provable, it would be right or correct [*richtig*] and that if its negation were provable, then $\mathrm{Bew}[R(q); q]$ would hold [*würde gelten*][2]. He then concludes this section by saying that "[f]rom the remark that $[R(q); q]$ says about itself that it is not provable, it follows at once that $[R(q); q]$ is true [*richtig ist*], for $[R(q); q]$ *is* indeed unprovable (being undecidable)" (Gödel *op. cit.*: [176], 151).[3] There is no mention of truth either in the formulation of Theorem VI, which claims that:

> GÖDEL'S ([1931], 1986a) THEOREM VI
> For every ω-consistent recursive class k of FORMULAS there are recursive CLASS SIGNS r such that neither $v\,\mathrm{Gen}\,r$ nor $\mathrm{Neg}(v\,\mathrm{Gen}\,r)$ belongs to $\mathrm{Flg}(k)$ (where v is the FREE VARIABLE of r).

But although Gödel's own formulations do not involve a direct or explicit claim to the effect that the undecidable formula is true, it can hardly be maintained that the formula which is undecidable *modulo* the consistency and ω-consistency of P, and which states that it is neither provable nor refutable in the system, may not be a bearer of truth.

As far as the informal presentation is concerned, the undecidable formula says *truly* of itself that it is not provable for it *is* indeed not provable. Is the case different when, instead of referring to the undecidable formula by means of its metamathematical description $[R(q); q]$, we refer to it my means of its so-called 'Gödel number' once we have determined the number q, i.e. by the expression '$17\,\mathrm{Gen}\,r$' ('$x\,\mathrm{Gen}\,y$' denoting the 15^{th} number theoretic function proven to be recursive)? It is true that Gödel ([1931: 189], 1986a: 177) concludes his proof of the first incompleteness theorem by saying "$17\,\mathrm{Gen}\,r$ is therefore undecidable on the basis of k, which proves Theorem VI" and *not* by saying "$17\,\mathrm{Gen}\,r$ is therefore true and undecidable on the basis of k, which proves Theorem VI". But the undecidable formula nevertheless *truly* says that $17\,\mathrm{Gen}\,r$ is not k-PROVABLE and that $\mathrm{Neg}(17\,\mathrm{Gen}\,r)$ is not k-PROVABLE.

[2] *Würde gelten* may also be rendered as "would be valid" or "would have authority".

[3] Note that *richtig* could be translated by *correct* instead of *true* and that Gödel does not use the german *wahr* in this instance. It is Jean van Heijenoort who uses the english equivalent of that german word in his translation. Kleene, in his presentation, also claims that $\mathscr{A}$ is "unprovable, hence true" (Kleene 1986: 128).

In both cases, that of '$[R(q); q]$' and that of '$17\,\mathrm{Gen}\,r$', we may construct an individual quotation-mark name of the formula *à la* Tarski, so as to get a specific instance of Tarski's schema "'p' is true if and only if p", "p is true" being a mere metalinguistic variant of p and the truth-conditions of the formula being mere redundant truth-conditions. Nothing requires here that truth be a substantial property over and above disquotational truth and so our predication of truth conforms to the minimalist view advocated by Paul Horwich (1990), according to which the predicate 'is true' is not used to attribute a genuine property. The predication of truth to the Gödel formula '$17\,\mathrm{Gen}\,r$' yields an expression, namely "'$17\,\mathrm{Gen}\,r$' is true", which is strictly equivalent to an expression which contains no mention of truth, namely '$17\,\mathrm{Gen}\,r$' itself.

Bearing this in mind, let me now sketch an outline of the semantic formulation of Gödel's proof.

Gödel's first incompleteness theorem exhibits an elementary formula, finitary in Hilbert's sense, which is proven to be both unprovable and irrefutable in P in the following way, as a result of the following steps :

(1) A formula $\mathscr{A}$ is constructed by diagonalization, which asserts its own unprovability in P.

(2a) The consistency of P being taken as a hypothesis, it is proven that $\mathscr{A}$ is unprovable in P.

(2b) We may then conclude that $\mathscr{A}$ is true since it asserts its own unprovability in P.

(3) The ω-consistency of P being taken as a hypothesis, it is proven that $\mathscr{A}$ is irrefutable in P.

(4) We may then conclude that $\mathscr{A}$ is undecidable in P.

(5) We may then conclude that $\mathscr{A}$ is true and that $\mathscr{A}$ is undecidable in P.

Our warrant for this last conclusion is that Gödel's formula $\mathscr{A}$ is both recognized to be true (step (2b)) *and* proven to be undecidable (step (4)), since it is both proven to be unprovable (step (2a)) and proven to be irrefutable (step (3)).

I now want to ask questions about three doctrines which are part of realism as described by Dummett[4], in relation to the Gödel formula. The first doctrine is that our understanding of the meaning of a statement amounts to a knowledge of its truth-conditions. Does our understanding of the meaning

[4] See, in particular, the articles "Truth" and "The Philosophical Basis of Intuitionistic Logic", which appear, respectively, as chapters 1 and 14 of Dummett (1978).

of $\mathscr{A}$ amount to such a knowledge? The second is that, if a statement is true, there must be something in virtue of which it is true. I have claimed that $\mathscr{A}$ is true; but in virtue of what is it true? Thirdly and crucially, for that is the defining thesis of realism, there is the idea that some truths may transcend all possible verification? Is $\mathscr{A}$ such a truth?

First of all, do we know the truth-conditions of $\mathscr{A}$ and can we manifest that knowledge? We must answer this question in the positive since the truth of $\mathscr{A}$ is *recognized* at step (2b).[5] We do, in effect, have a means at our disposal to find out that these truth-conditions are satisfied and we make the knowledge manifest by proving the unprovability of $\mathscr{A}$ under the hypothesis that P is consistent (step (2a)) and by concluding that $\mathscr{A}$ is true (step (2b)). Our answer is positive because, as Gödel ([1934: 21], 1986c: 362–363) points out:

> we can construct propositions which make statements about themselves [...]. It is even possible, for any metamathematical property f which can be expressed in the system, to construct a proposition which says of itself that it has this property.

If it therefore possible, for any predicate F of the language L_P of P expressing in P a given metamathematical property, to construct by diagonalization a formula $\mathscr{A}$ of L_P which asserts of itself that it possesses that property. If we note the Gödel number of that formula with the symbol '$\langle\mathscr{A}\rangle$', then, for every predicate F of L_P, there exists a formula such that $\mathscr{A} \iff F(\langle\mathscr{A}\rangle)$.

Let us choose as a metamathematical property the property of non-provability in P, expressed in P by the predicate 'non-Pr_P'. We then may construct a formula $\mathscr{A}$ which asserts its own unprovability, such that: $\mathscr{A} \iff \text{non-}\mathrm{Pr}_P(\langle\mathscr{A}\rangle)$.

Once step (1) is accomplished, we may then proceed to step (2a), and distinguish the following sub-steps leading to (2b).

If $\mathscr{A}$ were provable in P, then:

(2a1) $\mathrm{Pr}_P(\langle\mathscr{A}\rangle)$ would be true in P and therefore provable in P and

(2a2) non-$\mathrm{Pr}_P(\langle\mathscr{A}\rangle)$ would be provable in P, since $\mathscr{A}$ and non-$\mathrm{Pr}_P(\langle\mathscr{A}\rangle)$ are equivalent.

[5] This claim must be qualified. At steps (2a)–(2b), we manifest our knowledge that, *if* P is consistent, then $\mathscr{A}$ is unprovable in P and therefore true. There remains the further problem of knowing how we could know that P is consistent and make that knowledge manifest. I have focused here on the consequent of the conditional, but it is obvious, in view of Gödel's second incompleteness theorem — Theorem XI in Gödel ([1931]) — that a proof of the consistency of P cannot be obtained in P. What we have here, therefore, is a partial manifestation of our knowledge of the truth-conditions of $\mathscr{A}$.

(2a3) P would therefore be inconsistent.

(2a4) Under the hypothesis that P is consistent, $\mathscr{A}$ is therefore unprovable in P.

As Gödel himself noted ([1931: 176n15], 1986a: 151n15):

> Contrary to appearances, such a proposition [which says about itself that it is not provable] involves no faulty circularity, for initially it [only] asserts that a certain well-defined formula (namely, the one obtained from the qth formula in the lexicographic order by a certain substitution) is unprovable. Only subsequently (and so to speak by chance) does it turn out that this formula is precisely the one by which the proposition itself was expressed.

We may then directly proceed to step (2b): since $\mathscr{A} \iff \text{non-Pr}_P(\langle\mathscr{A}\rangle)$, $\mathscr{A}$ is true.

The question of determining whether or not we know the truth-conditions of $\mathscr{A}$ may be answered positively by the time we reach step (2b), for we make it plain, by proceeding from step (2a1) to step (2b), that we know that these truth-conditions are satisfied. So it is perfectly possible for us to manifest a knowledge of the truth-conditions of a formula proven to be unprovable in a certain formal system.

We may conclude from this that we are not here in a situation in which we could have a reason — of an anti-realist kind, in Dummett's sense of that expression — to eschew the notions of truth and truth-conditions altogether, or to replace truth by justification and truth-conditions by conditions of justification, on the basis that we would not be capable of recognizing whether or not conditions of the first kind obtain[6], for we are indeed, here, capable of such a recognition.

Moreover, since, as Gödel ([1934: 21], 1986c: 362) claimed, arithmetic propositions which make statements about themselves and involve only recursively defined functions are "undoubtedly meaningful statements", there is no reason to believe that the truth-conditional principle may not apply to $\mathscr{A}$. Its meaning *is* constituted by its truth-conditions and, accordingly, our knowledge of its meaning amounts to a knowledge of its truth-conditions. This answers the question related to the second doctrine: there is indeed something in virtue of which $\mathscr{A}$ is true, namely the proof which proceeds from step (1) to step (2b).

What about the third doctrine, i.e. the defining thesis of realism? May we draw a positive conclusion from the proof that some elementary formula is

[6] The argument for the replacement is found, e.g. in Dummett (1978: 225–227).

both unprovable and irrefutable? Gödel's proof is categorical on the following point: the elementary formula $\mathscr{A}$, proven to be undecidable in P given the consistency and ω-consistency of P, is thereby true in a sense of 'true' which may not offend a constructivist or anti-realist in Dummett's sense, for the formula is *recognized* to be true.

The proof of the first incompleteness theorem does not show that the truth-conditions of $\mathscr{A}$ transcend its conditions of justification. It shows something quite different, namely, as Dubucs (1991: 57) points out, that:

> our capacities for justification go beyond what is strictly speaking provable in a formal system: there exists, for each formal system which is sufficiently rich [i.e. such that the property 'provable in the system' is expressible in the system], undecidable elementary statements which we nevertheless have cogent reasons to hold as true.

In other words, Gödel's first incompleteness theorem does not show that the extension of the predicate 'true' is larger than the extension of the predicate 'recognizable as true'. Such a gap between truth and its recognition would indeed be unacceptable from a constructivist or anti-realist standpoint. What the theorem shows is that the extension of the predicate 'recognizable as true' exceeds the extension of the predicate 'provable in P', which is a quite different matter. It shows that the extensions of these two predicates do not coincide. What we have acknowledged so far is that the truth-conditions of $\mathscr{A}$ are transcendent with respect to its provability in P. But this hardly shows that we are in a situation where our acceptance of the truth-conditional principle should be judged problematic by an anti-realist, on the ground it would lead us to the admission of the possibility of some recognition-transcendent truth. On the contrary, it is precisely because $\mathscr{A}$ is not provable in P and, a fortiori, because it is proven not to be provable in P, that the endorsement of the truth-conditional principle is perfectly legitimate in this instance. By carrying out steps (2a1) to (2b), we make clear that we know the truth-conditions of $\mathscr{A}$ and these, by the nature of the case, are not recognition-transcendent.

We may now draw two conclusions.[7] First of all, there is no elementary formula whose truth could be undetectable in a given formal system if we assume that system to be consistent. The most we can say is that there

[7] To be precise, the conclusions are drawn from the first part of Gödel's proof and are grounded on the steps of its semantic formulation up to step (2b). I have not taken into account the proof of the unprovability of $\mathscr{A}$ (given the ω-consistency of P). The same conclusions would a fortiori be justified if the complete proof were taken into consideration.

are elementary formulae whose truth is algorithmically undetectable given the consistency of the system, which is a quite different thing. This is precisely the case with $\mathscr{A}$. No algorithmic procedure may help us to conclude that it is true, but its truth, although algorithmically undetectable in P, is nevertheless detectable by a *reductio ad absurdum* of the supposition that it is provable in P, given that P is consistent. So, unless we decide that the unavailability of an algorithmic procedure for deciding the truth-value of a formula is a criterion of the undetectability of its truth, there are no undetectable truths in a formal system if that system is consistent.

Should such an unavailability be a criterion? Step (2b) suggests the contrary. The second lesson we may draw is that we must distinguish the case of algorithmic undecidability from the case of undetectability of truth-value *simpliciter*, when discussing the question of knowing whether the modal schema '$\Diamond$(p is true $\cdot$ p is undecidable)' and its various instances is legitimate.

We may not merely oppose the undetectability or unrecognizability of truth (or, better, the possibility of its undetectability or unrecognizability) to algorithmic provability in P. We need to take into account a finer-grained distinction.

We must distinguish between:

(1) The gap between what is true in the standard model for arithmetic and what is recognizable as true on the basis of cogent reasons

and

(2) The gap between what is recognizable as true on the basis of cogent reasons and what is algorithmically recognizable as true in the standard model for arithmetic.

The question I wanted to consider was whether or not Gödel's first incompleteness theorem could help us to choose between intuitionist and classical semantics by showing that there could be undetectable arithmetical truths. The answer is clearly no, for the first gap is filled by the proof of the unprovability of $\mathscr{A}$ and, a fortiori, by the (complete) proof of its undecidability. The second, which may not be filled, just because of the same proof, leaves the debate about the legitimacy of the modal schema entirely open.

Acknowledgements. I have benefited from criticisms and comments by Wilfrid Hodges and Jan Woleński when I gave a shortened version of the paper at the Second German-Polish Workshop on Logic and Logical Philosophy in Żagań in April 1998. Other comments and criticisms were voiced by Karl Georg Niebergall and Matthias Schirn when I gave a somewhat longer version at the Institut für Philosophie, Logik und Wissenschaftstheorie of the University of Munich in June 1998. I have also benefited from the discussion which followed.

References

Boolos, G. , 1989, "A new proof of the Gödel incompleteness theorem", Notices of the American Mathematical Society, vol. 36, n° 4, p. 388–390.

Brouwer, L. E. J., [1948] 1983, "Consciousness, philosophy and mathematics", in *Philosophy of Mathematics – Selected Readings*, P. Benacerraf and H. Putnam (eds.), Cambridge U. P., 2nd edition, p. 90–96.

Dubucs, J., 1991, "La philosophie de Kurt Gödel [Compte rendu de lecture de *Reflections on Kurt Gödel* de Hao Wang]", *Philosophie de la logique et philosophie du langage I*, L'Age de la Science – Lectures Philosophiques, vol. 4, éd., J. Bouveresse, Odile Jacob, Paris, p. 53–68.

Dummett, M., A. E., 1978, *Truth and Other Enigmas*, Harvard U. P., Cambridge, Mass.

Glivenko, V. I., 1929, "Sur quelques points de la logique de M. Brouwer", Académie royale de Belgique, Bulletin de la classe des sciences (5), 15, p. 183–188.

Gödel, K., [1931] 1986a, "Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I" / "On formally undecidable propositions of *Principia Mathematica* and related systems I", *Collected Works*, vol. I: *Publications 1929–1936*, S. Feferman, ed.-in-chief, Oxford U. P., Oxford, trans. by J. van Heijenoort, p. [173–198] 144–195.

— [1933] 1986b, "Zur intuitionistischen Arithmetic und Zahlentheorie" / "On intuitionistic arithmetic and number theory", *Collected Works*, vol. I: *Publications 1929–1936*, S. Feferman, ed.-in-chief, Oxford U.P., Oxford, trans. by S. Bauer-Mengelberg and J. van Heijenoort, p. [34–38] 287–295.

— [1934] 1986c, "On undecidable propositions of formal mathematical systems", *Collected Works*, vol. I: *Publications 1929–1936*, S. Feferman, ed.-in-chief, Oxford U. P., Oxford, p. [1–27] 346–371.

— [1941] 1995, "In what sense is intuitionistic logic constructive?", *Collected Works*, vol. III: *Unpublished Essays and Lectures*, S. Feferman, ed.-in-chief, Oxford U. P., Oxford, p. [1–30] 189–200.

Horwich, P., 1990, *Truth*, Basil Blackwell, Oxford.

Kleene, S. C., 1986, "Introductory note to *1930b*, *1931* and *1932b*", Kurt Gödel: *Collected Works*, vol. I: *Publications 1929–1936*, S. Feferman, ed.-in-chief, Oxford U. P., Oxford, p. 126–141.

Shanker, S. G., 1989, "Wittgenstein's remarks on the significance of Gödel's theorem", in *Gödel's Theorem in Focus*, S. G. Shanker (ed.), Routledge and Kegan Paul, London and New York, p. 155–256.

FABRICE PATAUT
Institut d'Histoire et Philosophie des Sciences et des Techniques
C.N.R.S., Université de Paris 1
Paris, France
e-mail: pataut@idf.ext.jussieu.fr