Massimiliano Carrara
Enrico Martino

A Note on Gödel, Priest and Naïve Proof

Abstract. In the 1951 Gibbs lecture, Gödel asserted his famous *dichotomy*, where the notion of *informal proof* is at work. G. Priest developed an argument, grounded on the notion of naïve proof, to the effect that Gödel's first incompleteness theorem suggests the presence of *dialetheias*. In this paper, we adopt a plausible ideal notion of naïve proof, in agreement with Gödel's conception, superseding the criticisms against the usual notion of naïve proof used by *real working mathematicians*. We explore the connection between Gödel's theorem and naïve proof so understood, both from a classical and a dialetheic perspective.

Keywords: incompleteness; standard model; naïve proof; dialetheia; liar paradox; Curry's paradox

1. Introduction

In the 1951 Gibbs lecture, Gödel asserted his famous *dichotomy* that establishes a widely discussed connection between his incompleteness theorems and the notion of *informal proof*. Such a notion has been here and there discussed in philosophical and logical literature. Specifically, in "The logic of paradox" (Priest, 1979) developed an argument, grounded on the notion of naïve proof, to the effect that Gödel's first incompleteness theorem suggests the presence of *dialetheias* (viz., sentences that are both true and false) in the standard model of arithmetic. This last point has been amplified in his "Is arithmetic consistent?" (Priest, 1994). Although Chihara (1984), Shapiro (2002), Berto (2009) and others criticized the argument, Priest proposed it again in the first

edition of *In Contradiction* and amplified it in the second edition (Priest, 2006) (discussed by Tanswell, 2016, among the others).

Much of the criticism was directed against the notion of naïve proof itself, particularly against the thesis that everything that is naïvely provable is true. If the notion of naïve proof is understood as embracing all proofs performed by real working mathematicians, as Priest seemed to suggest (e.g. Priest, 1994), the thesis is surely untenable. However, here, we adopt a plausible *ideal* notion of naïve proof, in agreement with Gödel's conception of informal proof. We think that it is worthwhile to explore here the connection between Gödel's theorem and naïve proof, even from a dialetheic perspective, where the consistency of naïve provability is not assumed.[1]

2. On naïve proof

Priest characterises the notion of naïve proof (n-proof) as follows:

> Proof, as understood by mathematicians (not logicians), is that process of deductive argumentation by which I establish certain mathematical claims to be true. In other words, suppose I have a mathematical assertion, say a claim of number theory, whose truth or falsity I wish to establish. I look for a proof or a refutation, that is a proof of its negation... I will call the informal deductive arguments from basic statements naïve proofs.   (Priest, 2006, p. 40)

According to Priest, an alleged paradox is suggested by the analogy drawn between the familiar informal proof of Gödel's undecidable sentence G and the liar paradox:

> As is clear to anyone who is familiar with *Gödel's* theorem, at its heart there lies a paradox. Informally the 'undecidable' sentence is the sentence 'this sentence is not provable'. Suppose that it is provable; then, since whatever is provable is true, it is not provable. Hence it is not provable. But we have just proved this. So it is provable after all (as well).   (Priest, 2006, p. 237)

The paradox at issue arises in the natural language from the liar paradox by replacing the notion of truth with that of informal provability. Call it the *unprovability paradox*. Like the liar paradox, it is a typical paradox

---

[1] On the same topic of this paper see Shapiro's interesting recent paper (2019), specifically sec. 22.5 (entitled "Gödel Meets Curry"), and Priest's reply to Shapiro (Priest, 2019), especially sec. 27.21.3 on "Curried Undecidability".

of self-reference generated by the sentence

(*) This sentence is unprovable.

One can reason on (*) as follows:

(*) is provable or unprovable. If it is provable, then it is true and hence unprovable. Thus, in any case, it is unprovable and hence true. So, it is unprovable. But we have just proved it. So, it is false. Thus, classically, (*) is paradoxical; dialetheically, it is a dialetheia.

It is evident that the unprovability paradox inspired Gödel's incompleteness theorem. It is important, however, for our purposes, to stress that no self-reference is involved in the latter. Gödel's unprovable sentence G is purely arithmetical; that is, it is built up in terms of the primitive arithmetical notions. As with any arithmetical sentence, it involves neither the notion of provability nor any sort of self-reference.

To be more explicit, recall the arithmetical translation of the notion of formal provability in Peano arithmetic (PA).

Let L be the language of first-order arithmetic and S any sound extension of PA. Soundness is here understood in the sense that all S-provable sentences are true in the standard model N of arithmetic. The metalinguistic relation that correlates a sentence with a (possible) S-proof of it is expressible arithmetically through a decidable formula $\mathrm{Prf}(x, y)$ (read: y is a proof of x), such that:

(i) if n is the code of a proof of ϕ then $\mathrm{Prf}(\langle\phi\rangle, n)$ is PA-provable (where $\langle\phi\rangle$ is the code of ϕ), and

(ii) if n is not the code of an S-proof of ϕ, then $\neg\mathrm{Prf}(\langle\phi\rangle, n)$ is PA-provable.

The arithmetical translation of S-provability is the predicate $P(x) =_{\mathrm{df}} \exists y \mathrm{Prf}(x, y)$. Therefore, for any sentence A, $P(\langle A\rangle)$ is true iff it is S-provable. Let $G = \neg P(\langle G\rangle)$ be the Gödel sentence relative to S. Thus, G is true iff it is not S-provable.

On one hand it turns out that G is S-provable iff $\neg G$ is S-provable. Since S is sound, it follows classically that G is S-undecidable. On the other hand, a dialetheist may conjecture that both G and $\neg G$ are S-provable so that G is a dialetheia. However, observe that such a conjecture cannot be supported by the unprovability paradox.

In contrast with the ungrounded (in Kripke's sense) sentence (*), the truth conditions of G are well defined in arithmetical terms. The fact

that we can prove, in the meta-language, that it is true if, and only if, it is unprovable, has nothing to do with its meaning. It does not say of itself to be S-unprovable; it says that the number $\langle G \rangle$ does not satisfy the arithmetical predicate $P(x)$.

Both the classicist and the dialetheist can infer G from the soundness of S by the following:

Simple Proof. G is S-provable or not. If it is S-provable, then $\neg P(\langle G \rangle)$ is true, that is, it is S-unprovable.Thus, in any case, it is unprovable. Hence, it is true.

Observe that this proof is acceptable even by a dialetheist, since it does not exploit the consistency but only the soundness of S.

3. Gödel's dichotomy

Let us recall Gödel's dichotomy:

> Either mathematics is incompletable in this sense, that its evident ax-ioms can never be comprised in a finite rule, that is to say, the human mind (even in the realm of pure mathematics) infinitely surpasses the powers of any finite machine, or else there exist absolute unsolvable Diophantine problems of the type specified.   (Gödel, 1951, p. 310)

Gödel provides a further explanation of his disjunction:

> [The incompleteness theorem] makes the incompletability of mathemat-ics particularly evident. For it makes it impossible that someone should set up a certain well-defined system of axioms and rules and consistently make the following assertion about it: all of these axioms and rules I perceive (with mathematical certitude) to be correct, and moreover I believe that they contain all of mathematics. If someone makes such a statement, he contradicts himself. For if he perceives the axioms under consideration as to be correct, he also perceives (with the same cer-tainty) that they are consistent. Hence, he has a mathematical insight not derivable from his axioms.   (Gödel, 1951, p. 310)

Though Gödel correlates correctness with consistency, his argument holds even dialetheically by understanding correctness as soundness in the above sense.

We think that by "the human mind", Gödel means the mind of an idealised human being, free of empirical space-time limitations. To make

the idea more explicit, let us introduce a fictional character, Hans. He is thought of as a classical mathematician, whose reasoning about the standard model N of arithmetic is free of errors and of empirical space-time limitations. We say that an arithmetical sentence is *naïvely provable* (briefly, *n*-provable) if Hans can informally recognise its truth in N with *absolute mathematical certainty*. As an alleged candidate for formalising *n*-provability, Gödel means an axiomatic system with correct axioms and inference rules. Tanswell (2016) discusses at length the difficulties of translating *informal reasoning* into a unique formal system. In the present paper we don't want to enter in such discussion.

By a formalisation of *n*-provability we will simply mean a recursively enumerable (briefly r.e.) predicate $P(x)$ of first order PA-arithmetic, satisfied by all and only all (the codes of) *n*-provable arithmetical sentences. We don't make any assumption on the structure of *n*-proofs; in particular, we don't assume that they are expressible in the language L of first order arithmetic. According to this understanding, no formalization of the language of *n*-proofs is presupposed by a formalisation of *n*-provability.

Thus, if P formalises *n*-provability, given any P-sentence A (i.e. such that $P(\langle A \rangle)$), Hans can *n*-prove A. However, this does not entail that Hans can *n*-prove that P is sound, that is, that all infinitely many P-sentences are true. It may happen that he can recognise the truth of the P-sentences only one at a time.

Given any r.e. predicate $P(x)$, if Hans can *n*-prove P-soundness, then, by using the simple proof, he can *n*-prove Gödel's P-unprovable sentence G_P. In this sense, we can say that his mind surpasses the power of P. In this case, P fails to formalise *n*-provability. On the other hand, if P formalizes *n*-provability, Hans cannot *n*-prove P-soundness. In this case G_P is true but, classically, absolutely unprovable, i.e. even *n*-unprovable. So, we can reformulate Gödel's dichotomy as follows:

Gödel's dichotomy. Either no arithmetical predicate P can represent all *n*-provable sentences or there is an absolutely unprovable arithmetical sentence.

Several authors (classical *loci* are Lucas (1961), Penrose (1994)) have argued for the first horn, but no one, as far as we know, has produced a convincing argument that has settled the matter. Anyway, we think that the first horn is highly plausible, because the creative resources of the human mind for inventing an informal proof do not seem to be a priori

determined. Furthermore, an informal proof of an arithmetical sentence can exploit our basic intuition of *absolute finiteness* (i.e. independent of a set-theoretical model) that characterizes the standard model N of arithmetic, where every number follows finitely many numbers. In contrast, no formal system can capture such an intuition, as the presence of non-standard models shows.

Remember Hilbert's conviction of the solvability of all mathematical problems, expressed as an axiom in his lecture at the international congress in Paris, 1900:

> Is the axiom of solvability of every problem a peculiar characteristic of mathematical thought alone, or is it possibly a general law inherent in the nature of the mind, that all questions which it asks must be answerable? This conviction of the solvability of every mathematical problem is a powerful incentive to the worker. We hear within us the perpetual call: There is the problem. Seek its solution. You can find it by pure reason, for in mathematics there is no ignorabimus.
>
> (Hilbert, 1900, as translated in Browder Hilbert, 1976, p. 7)

Furthermore, observe that, since P is r.e., the absolutely unprovable sentence G_P would be a Π_1-sentence, that is of form $\forall x R(x)$, where $R(x)$ is a decidable predicate. It is hard to imagine what could deny Hans *a priori* the possibility of recognizing, by means of a suitable reflection on the deciding rule, that this yields, for every input, the output *true*.

Be that as it may, from Gödel's dichotomy, this proposition immediately follows:

(2.1) If P formalises n-provability, then Hans cannot recognise the P-soundness.

As we know, from P-soundness, the truth of G_P follows. Thus, if Hans could recognise P-soundness, he could obtain an n-proof of G_P, and hence the P-unprovability of G_P, against the hypothesis that P represents n-provability.

Now, we want to emphasise that the formalisability of n-provability may be understood according to two different ways:

1-*formalisability.* Hans can find a suitable arithmetical predicate P and n-prove that it codifies n-provability.

2-*formalisability.* There exists *factually* an arithmetical predicate P that codifies n-provability (independently of the possibility of knowing this fact).

Therefore, we can rephrase (2.1) as follows:

(2.2) n-provability fails to be 1-formalisable.

It is worth noting, in passing, that according to the intuitionistic perspective, the two meanings of formalisability collapse. The reason is that the mere existence of an arithmetical predicate satisfying certain unknowable conditions is, for an intuitionist, unintelligible. Thus, for an intuitionist, the first horn of Gödel's dichotomy holds, while the second does not, since there is no room, intuitionistically, for true unprovable sentences. (See on that Martin-Löf, 1995).

However, we think that, when dealing with an attempt to formalise naïve proof, even for a classicist (as well as for a dialetheist), the more interesting notion is 1-formalisability. If formalisation is perceived as a device for analysing informal proof, the mere existence of an unknowable system that *in fact* proves all n-provable sentences is of no help for the purpose.

What about Gödel's dichotomy from the dialetheic perspective? Here, there is room for a third horn:

Dialetheic trichotomy. (i) No arithmetical predicate represents n-provability, or (ii) some arithmetical sentence is absolutely unprovable, or (iii) some arithmetical sentence is a dialetheia.

Let us introduce an ideal dialetheist, Otto, free of empirical limitations, who reasons, without errors, in accordance with dialetheic logic. (2.1) becomes

(2.3) If P represents n-provability, then either Otto cannot recognise P-soundness, or G_P is a dialetheia.

(2.2) becomes

(2.4) If n-provability is 1-formalisable then some sentence is a dialetheia.

Priest (for example in 1994) does not consider Gödel's dichotomy nor the dialetheic trichotomy. He tries to argue directly for the existence of a dialetheia as a consequence of Gödel's first incompleteness theorem. We think, however, that the above trichotomy can be useful for the discussion of his argument.

Remember that, in order to recover the notion of exclusivity (banished from logical negation), Priest introduces the notion of *rejection*, according to which one cannot accept and reject the same proposition.

Using such a notion, a dialetheist can immediately reject the existence of arithmetical dialetheias. As already observed, the truth conditions of any arithmetical sentence A are well defined in terms of the primitive arithmetical notions. Now, if A is atomic, its truth value is determined by a computation that yields a unique value (remember that Otto is free of empirical limitations, so that he is not affected by the empirical difficulties of performing any long computation). Thus, the assumption that A is a dialetheia is clearly rejectable.

Then, the rejection can be forwardly extended to any arithmetical sentence by an induction on its complexity. We therefore think that a dialetheist should accept the following proposition:

(2.5) The existence of arithmetical dialetheias is rejectable.

Thus, if Priest's argument were conclusive, we would obtain an authentic dialetheic paradox. Let us examine Priest's argument.

4. Priest's argument

Priest extends the language L of first-order arithmetic with a predicate $\beta_\mathsf{N}(x)$ on natural numbers, which expresses the fact that x is (the code of) an n-provable sentence. He claims that the predicate satisfies the following principles:

(1) $\vdash_\mathsf{N} \beta_\mathsf{N}\langle\alpha\rangle \supset \alpha$,
(2) If $\vdash_\mathsf{N} \alpha$ then $\vdash_\mathsf{N} \beta_\mathsf{N}\langle\alpha\rangle$,

where $\vdash_\mathsf{N}$ expresses naïve provability. These principles are justified as follows:

> For (1), it is analytic that whatever is naïvely provable is true. [...] And since this is analytic it is itself naïvely provable. For (2), if something is naïvely proved then this fact itself constitutes a proof that α is provable.
> (Priest, 2006, p. 238)

Then, by the usual method of self-reference, Priest constructs a sentence γ of form $\neg\beta_\mathsf{N}\langle\gamma\rangle$ and uses the simple proof to show that both γ and $\neg\gamma$ are provable. He concludes that arithmetic is inconsistent. Priest explicitly declares that he has not assumed that β_N itself is a predicate that can be constructed from the usual arithmetic vocabulary $(', +, \times)$.

Now, on one hand, if β_N is not arithmetical but is *intensionally* introduced as the predicate of naïve provability of the arithmetical sentences,

principles (1) and (2) are to be restricted to any *arithmetical* α. And since γ is not arithmetical no contradiction follows. On the other hand, if β_{N} is intensionally introduced as the predicate of naïve provability of the sentences of the *extended* language $\mathsf{L} \cup \{\beta_{\mathsf{N}}\}$, we get an authentic paradox of self-reference. In this case what is inconsistent is the *extended* theory. So, in either case the inconsistence of arithmetic does not follow.

Reminding the source of his interest in what he calls "l'*affaire* Gödel", Priest says:

> It seemed to me that the essence of the matter was how we show that the undecidable sentence is true in the standard model of arithmetic.[...] To carry out such reasoning, we need a language with a true predicate. And of course, if the undecidable sentence is indeed provable, the theory is inconsistent. So we have a contradiction on our hands. Paraconsistency was therefore required. (Priest, 2019, p. 648)

Observe, however, that the truth predicate needed in the naïve proof of the undecidable sentence is a metalinguistic (non arithmetical) predicate applicable only to arithmetical sentences. Paraconsistency was therefore not required.

Priest's claim that arithmetic is inconsistent seems to be nothing but a loose way of saying that the extended theory, with the new provability predicate and the new principles (1) and (2) for the extended language, is inconsistent.

It is worth noting, however, that, if β_{N} is understood as the n-predicate for the *extended* language, it is threatened by Curry's paradox. In some papers (for example in Beall and Murzi, 2013; Carrara et al., 2010; Carrara and Martino, 2011; Shapiro, 2011; Whittle, 2004; for a survey see Murzi and Carrara, 2015b, sec. 2.2) it is shown how, using logical principles in agreement with the *logic of paradox* (Priest (1979)), a Curry-like argument leads to the triviality of the extended theory. On the other hand, such conclusion can be avoided by using a suitable substructural logic (for an introduction see Murzi and Carrara, 2015b, sec. 2.3; and Ripley, 2015b).[2]

---

[2] Substructural approaches are sometimes met with skepticism. Field, for example, in his *Saving Truth from Paradox* says that he has not seen: "sufficient reason to explore this kind of approach (which I find very hard to get my head around), since I believe we can do quite well without it" (Field, 2008, pp. 10–11). However, as Murzi and Carrara observed: "While the paradoxes of naïve logical properties don't help one getting one's head around substructural consequence relations [...], these paradoxes put pressure on Field's claim that a substructural revision of classical logic

Be that as it may, subsequently, Priest suggests that β_{N} should be arithmetically expressible (Priest attributes this observation to A. Visser in (2006, p. 239, footnote 16)). Briefly, the suggestion arises from the fact — stressed by Dummett — that when a proof is presented to us, we are able — at least, in principle, — to recognise it as such:

> Intuitionists incline to write as though, while we cannot delimit in advance the realm of all possible intuitionistically valid proofs, still we can be certain for particular proofs given, and particular principles of proof enunciated, that they are intuitionistically correct.
>
> <div align="right">(Dummett, 1978, p. 347)</div>

Thus, assuming a numerical codification of the language of naïve proofs, the suggestion is that the set of (the codes of) naïve proofs should be decidable and hence, according to Church's thesis, recursive, so that the set of naïvely provable sentences should be recursively enumerable.

On the ground of these considerations, Priest maintains the existence of an aritmethical predicate β (without a subscript), coextensive with β_{N}. Gödel's corresponding sentence γ is arithmetical, and according to him, it would be a dialetheia. Thus, Priest concludes that arithmetic is contradictory. In the following section, we scrutinise Priest's argument in detail.

5. An analysis of Priest's argument

First, observe that Priest's appeal to Church's thesis for n-proofs seems to be deceptive. According to Church's thesis, any set of numbers, decidable by means of any *mechanical* procedure, is recursive. Now, we agree with Dummett that in principle, the set of naïve proofs must be decidable. This is to be understood in the sense that when an informal argument is given, an ideal mathematician is able to decide whether it is a correct proof or not, according to his own conception of informal but rigorous mathematical reasoning. Thus, Hans can decide if an alleged proof is classically correct; likewise, Otto is able to decide whether it is dialetheically correct. However, what is highly problematic is the question of whether such an ability is of a *mechanical* nature. Certainly,

---

is not needed in order to adequately deal with semantic paradoxes in general. Partly for this reason, the literature on substructural approaches, . . . , has recently been burgeoning." For references on substructural approaches see (Murzi and Carrara, 2015b, p. 10) and the footnote 3 below.

it is not for Dummett who, as a good intuitionist, explicitly endorses an anti-mechanistic conception of intuitive reasoning. In the previously quoted passage, he explicitly observes that "we cannot delimit in advance the realm of all possible intuitionistically valid proofs" (Dummett, 1978, p. 347).

On the contrary, the possibility of formalising the notion of n-proof presupposes that the set of n-proofs is a priori well determined. As already observed, this seems to be in contrast with the openness of mathematical creativity.[3]

Priest insists that a mechanistic conception of the human mind is needed for explaining how the notion of *naïve proof* is learned. We think, in contrast, that the notion of naïve proof is directly acquired from the understanding of the notion of truth, together with the notion of knowledge. Naïve proofs are of a semantic nature. Any argument that leads, with absolute certainty, to the knowledge of the truth of a sentence A counts as an n-proof of A. This seems to agree with Priest's above-quoted explanation of an n-proof as any deductive process by which one establishes that a mathematical claim is true. Besides, as already observed, an n-proof of an arithmetical sentence may use the fundamental intuition of the generative process of natural numbers, which no formal system can capture.

Anyway, let us assume, for the sake of discussion, that a suitable mechanistic philosophy of mind guarantees that n-provability is formalisable. Let us distinguish — as above — between 1-*formalisability* and 2-*formalisability*.

Following the recent literature on this topic (see, in particular, Beall and Murzi, 2013); on a survey on the topic with references (see Murzi and Carrara, 2015b, sec. 2.2) we show how also 1-*formalizability* is threatened by a Curry-like argument:

(4.1) If n-provability is 1-*formalisable*, then N is trivial.

Assume that Otto can find a r.e. predicate $P(x)$ and an n-proof that any arithmetical sentence A is n-provable iff $P(\langle A \rangle)$. In the same vein, we can assume that Otto can find a r.e. relation $R(x, y)$ and an n-proof

---

[3] A similar objection has been raised by Shapiro: "To invoke Church's thesis one must specify an algorithm, a step by step procedure for computing a value, or deciding a question, a procedure that involves no creativity or use of intuition. And that we clearly do not have. [...] An instruction to 'check whether such and such a string is a good proof' does not count as a step in an algorithm" (Shapiro, 2019, p. 472).

that for all sentences A and B, $R(\langle A \rangle, \langle B \rangle)$ iff B is n-deducible from A (this means that it is n-provable that from the truth of A in N, the truth of B in N follows). Given an arbitrary sentence C, Otto can find, by diagonalisation, a sentence H of form $R(\langle H \rangle, \langle C \rangle)$. Now, Otto, can reason as follows:

Assume that H is true. Then, $R(\langle H \rangle, \langle C \rangle)$ is true; hence, C is n-provable from H, and since H, by hypothesis, is true, so is C. Thus, I have proven C from H. Therefore, $R(\langle H \rangle, \langle C \rangle)$ is true, that is, H is true. Since C is n-provable from H, C is true. Since C is arbitrary, N is trivial.

This argument is clearly inspired by Curry's paradox and it is extensively discussed in the literature.[4] A number of controversial strategies — in particular substructural logics — of restricting the ordinary logic to block Curry's paradox have been proposed by several authors (for an introduction to this topic see (Ripley, 2015b)).[5]

So one may wonder if the argument we proposed, which rests on ordinary logic, is reliable. An answer may be that a naïve logic, insofar as it is compatible with the existence of dialetheias, is fitting to Otto's naïve reasoning. We think there is another possible answer: we are not dealing with Curry's paradox. As already emphasised, the latter is a paradox of self-reference, arising in languages where ungrounded sentences (in Kripke's sense) are allowed, whereas our Curry-like argument is applied to first-order arithmetical sentences.

It may be instructive to compare the Curry-like argument we propose with one, apparently similar. In his paper (2015b) Ripley gives a formulation of the paradox, using a validity predicate formally, similar to our $R(x, y)$. He extends the object language by means of a predicate $V(x, y)$, understood as a validity predicate for the extended lan-

---

[4] For a survey see, again, (Murzi and Carrara, 2015b). See also the special issue edited by Murzi and Carrara (2015a) on *Paradox and Logical Revision* for an extensive number of papers on this topic.

[5] Substructural approaches can be either of the *SContr-free* (from Γ, ϕ, $\phi \vdash \psi$ it follows that Γ, $\phi \vdash \psi$) or *Cut-free* (from $\Gamma \vdash \phi$ and Δ, $\phi \vdash \psi$ it follows that Γ, $\Delta \vdash \psi$). On the first approach, *SContr-free*, see (Fitch, 1942, 1948, 1950) and recently, among the others, (Caret and Weber, 2015; Mares and Paoli, 2014; Murzi and Shapiro, 2014; Priest, 2015; Shapiro, 2011, 2013; Weber, 2014; Zardini, 2011, 2012, 2013a,b, 2014). On the second approach, the *Cut-free approach*, (Smiley, 1957)) and (Weir, 2005) have been recently, and influentially, followed by P. Cobreros, P. Egré, D. Ripley, R. van Rooij in (2012; 2014). In particular, Ripley in various papers (for example in 2013; 2015a) has developed an *inferentialist* argument against Cut.

guage (read $V(\langle A\rangle, \langle B\rangle)$ as "B follows from A"). The diagonal sentence $\nu = V(\langle \nu\rangle, \langle \bot\rangle)$ leads to trivialism by the same argument used in the proof exposed in (4.1) (see on this, among the others, Beall and Murzi, 2013). Observe, however, that there is difference between the contexts of $V(x, y)$ and $R(x, y)$: while ν is a self-referential sentence, which says of itself to imply $\bot$, H is an ordinary arithmetical sentence. The mere fact that it follows, under the assumption at work, that if H is true, so is $\bot$, has nothing to do with the arithmetical meaning of H, which does not involve any sort of self-reference. In sec. 2 we made a similar remark concerning the distinction between the genuine unprovability paradox and Gödel's undecidable sentence G. The same difference holds for the connection between the liar paradox and Tarski's theorem that the truth predicate T for arithmetic is not arithmetical. Classical logic, which, when applied to the liar sentence, leads to a paradox, seems to be adequate to bring to absurdity the assumption that any arithmetical predicate is extensionally equivalent to T. Similarly, coming back to the proof given in (4.1), we think that the mere fact that ordinary logic is inadequate to solve Curry's paradox by no means hampers the use of a Curry-like argument within a framework where no self-reference is involved. After all, among the various philosophical conceptions about logical paradoxes, one may embrace the traditional view, going back to Stoic philosophy, that defends ordinary logic and regards self-referential sentences as pathological.

Let us formalise the proof of (4.1) in natural deduction, so that one can judge whether the principles involved are dialetheically acceptable.

Let $\Gamma \Rightarrow A$ express that A is naïve consequence of Γ. More explicitly, $\Gamma \Rightarrow A$ is to be read: there is a naïve argument from the premise that all sentences in Γ are true in N to the conclusion that A is true in N.

Inference rules:

$$\Gamma, A \Rightarrow A \quad \text{(Reflection1)}$$

$$\frac{\Gamma \Rightarrow A}{\Gamma \Rightarrow A} \text{ (Reflection2)}$$

$$\frac{\Gamma \Rightarrow A, \ \Delta \Rightarrow R(\langle A\rangle, \langle B\rangle)}{\Gamma, \ \Delta \Rightarrow B} \text{ (R} - \text{Elimination)}$$

$$\frac{\Gamma, A \Rightarrow B}{\Gamma \Rightarrow R(\langle A\rangle, \langle B\rangle)} \text{ (R-Introduction)}$$

A formal deduction of the arbitrary sentence C is the following:

$$
\cfrac{
  \cfrac{
    H \Rightarrow H \,\text{(Ref.1)} \qquad
    \cfrac{
      \cfrac{H \Rightarrow R(\langle H\rangle, \langle C\rangle)\,\text{(Ref.1)}}
            {\cfrac{H \Rightarrow C}{\Rightarrow R(\langle H\rangle, \langle C\rangle)}\,\text{R-intr}}
    }{}
  }{\Rightarrow H \,\text{(Ref.2)}} \,\text{R-el}
  \qquad\qquad \Rightarrow R(\langle H\rangle, \langle C\rangle)\ \text{(Ref.2)}
}{\Rightarrow C} \,\text{R-el}
$$

Insofar as the above proof is dialetheically acceptable, we can conclude that, even from the dialetheic perspective, n-provability fails to be 1-*formalisable*, on pain of triviality.

Observe that this result, far from being paradoxical, is plausible. Since the great variety of possible means to obtain an n-proof is, *a priori*, unsurveyable, it is hard to imagine how, given any r.e. set of n-provable sentences, one could exclude that some further sentence is n-provable.

It is worth noting, however, that, even if 4.1 is dialetheically acceptable, it does not reject Priest's conjecture. The latter does not need 1-*formalizability*.

Let again $P(x)$ be a r.e. predicate that formalizes n-provability. Suppose that Otto is given such $P(x)$ through its definition. In order to prove that G_P is a dialetheia, Otto does not need to know that $P(x)$ exhausts all n-provable sentences; he needs only to recognise that it is sound. For, in this case, he is able to n-prove G_P, so that G_P is true and hence $\neg P(G_P)$. But, since, as a matter of fact, $P(x)$ formalizes n-provability, $P(G_P)$ holds. Besides, as $P(x)$ is r.e., Otto can recognize that $P(G_P)$ and conclude that G_P is a dialetheia.

One may wonder if, in a similar situation, relative to n-deducibility, Otto can use the Curry-like argument, to n-prove that N is trivial. The answer is no.

For, suppose, as above, that $R(x, y)$ formalizes n-deducibility and that Otto knows that R is sound but does not know that it is exhaustive. Then, he cannot use the $\mathsf{R - Introduction}$ rule since, as observed above, this exploits the exhaustiveness of R. So he cannot reach the step 5 of the Curry-like proof.

To sum up, the *mere* existence of a r.e. predicate $P(x)$ providing all and only all n-provable sentences does not entail that the relative Gödel sentence is a dialetheia, because it leaves open the second horn of the trichotomy, according to which G_P would be true but absolutely unprovable. To reject this horn, Otto should be able to n-prove that $P(x)$ is sound. If he succeeded to doing that, he would reach a n-proof

that G_P is a dialetheia and, in view of 2.5, an interesting dialetheic paradox would follow.

6. Conclusion

In this paper, we have discussed the problem of formalising the notion of naïve proof for arithmetical sentences from the classical and the dialetheic perspectives. We have considered an ideal notion of naïve proof, which we think is in line with Gödel's conception of such a notion. To make it more explicit, we have introduced an ideal classical mathematician Hans, as well as an ideal dialetheic mathematician Otto, both free of empirical limitations of space and time; besides, their proofs are always free of errors. We have distinguished between two notions of formalisability. The first, 1-*formalisability*, requires the possibility, in principle, of knowing that a certain system formalises n-provability. The second, 2-*formalisability*, requires the mere existence of such a system. By virtue of Gödel's dichotomy, 1-*formalisability* fails classically, while, in a dialetheic perspective, is affected by a Curry-like argument.

Priest's conjecture that naïve proof is formalisable is suggested by a doubtful interpretation of Church's thesis, which seems to presuppose, a priori, a mechanistic conception of informal mathematical reasoning.

Anyway, 2-*formalisability* may hold classically, according to the second horn of Gödel's dichotomy, as well as dialetheically, according to the second or the third horn of the dialetheic trichotomy. It is, however, inadequate, in itself, to show Priest's conjecture of the inconsistency of N. To prove the conjecture, Otto should recognise the soundness of the alleged predicate $P(x)$ that formalises n-provability. In this case he would reach an n-proof that G_P is a dialetheia and, in view 2.5, a noteworthy dialetheic paradox would follow.

References

Beall, J., and J. Murzi, 2013, "Two flavors of Curry's paradox", *The Journal of Philosophy* 110 (3): 143–65. DOI: `10.5840/jphil2013110336`

Berto, F., 2009, "The Gödel paradox and Wittgenstein's reasons", *Philosophia Mathematica* 17: 208–219. DOI: `10.1093/philmat/nkp001`

Caret, C. R., and Z Weber, (2015, "A note on contraction-free logic for validity", *Topoi* 34 (1): 63–74. DOI: `10.1007/s11245-014-9241-z`

Carrara, M., S. Gaio and E. Martino, 2010, "Can Priest's dialetheism avoid trivialism?", pages 53–64 in M. Peliš and V. Punčochoř (eds.), *The Logica Yearbook*, College Publications: London.

Carrara, M., and E. Martino, 2011, "Curry's paradox. A new argument for trivialism", *Logic and Philosophy of Science* 9: 199–206.

Chihara, C. S., 1984, "Priest, the liar, and Gödel", *Journal of Philosophical Logic* 13 (2): 117–124. DOI: `10.1007/BF00453016`

Cobreros, P., P. Egré, D. Ripley and van R. Rooij, 2012, "Tolerant, classical, strict", *Journal of Philosophical Logic* 41 (2): 347–85. DOI: `10.1007/s10992-010-9165-z`

Cobreros, P., P. Égré, D. Ripley, and van R. Rooij, 2013, "Reaching transparent truth", *Mind* 122 (488): 841–866. DOI: `10.1093/mind/fzt110`

Dummett, M., 1978, *Truth and Other Enigmas*, Duckworth: London.

Field, H., 2008, *Saving Truth from Paradox*, Oxford University Press: Oxford.

Fitch, F., 1942, "A basic logic", *Journal of Symbolic Logic* 7: 105–14. DOI: `10.2307/2269291`

Fitch, F., 1948, "An extension of basic logic", *Journal of Symbolic Logic* 13: 95–106. DOI: `10.2307/2267330`

Fitch, F., 1950, "A further consistent extension of basic logic", *Journal of Symbolic Logic* 14: 209–18. DOI: `10.2307/2269228`

Gödel, K., 1951, "Some basic theorems in the foundations of mathematics and their implications", pages 304–323 in S. Feferman et. al. (eds.), *Collected Works*, volume III, Oxford University Press: Oxford.

Hilbert, D, 1976, "Mathematische probleme, vortrag, gehalten auf dem internationalen mathematiker kongress zu paris" (1900), pages 1–34 in F. Browder (ed.), *Proceedings of Symposia in Pure Mathematics*, volume 28, American Mathematical Society, Providence.

Lucas, J. R, 1961, "Minds, machines, and Goedel", *Philosophy* 36: 112–137.

Mares, E., and F. Paoli, 2014, "Logical consequence and the paradoxes", *Journal of Philosophical Logic* 43: 439–469. DOI: `10.1007/s10992-013-9268-4`

Martin-Löf, P., 1995, "Verificationism then and now", pages 187–196 in M. van der Schaar (ed.), *The Foundational Debate: Complexity and Constructivity in Mathematics and Physics*, "Vienna Circle Institute Yearbook 3", Kluwer: Dordrecht.

Murzi, J., and M. Carrara (eds.), 2015a, "Paradox and logical revision", *Topoi*, volume 34.

Murzi, J., and M. Carrara, 2015b, "Paradox and logical revision: A short introduction", *Topoi* 34: 7–14. DOI: `10.1007/s11245-014-9286-z`

Murzi, J., and L. Shapiro, 2014, "Validity and truth-preservation", pages 431–460 in T. Achourioti, H. Galinon, J. Martinez-Fernandez and F. Fujimoto (eds.), *Unifying the Philosophy of Truth*, Springer: Berlin. DOI: `10.1007/978-94-017-9673-6_22`

Penrose, R., 1994, *Shadows of the Mind*, Oxford University Press: Oxford.

Priest, G., 1979, "The logic of paradox", *Journal of Philosophical Logic* 8: 219–241. DOI: `10.1007/BF00258428`

Priest, G., 1994, "Is arithmetic consistent?", *Mind* 103: 337–349. DOI: `10.1093/mind/103.411.337`

Priest, G., 2006, *In Contradiction*, Oxford University Press: Oxford. Expanded edition (first published 1987, Kluwer: Dordrecht).

Priest, G., 2015, "Fusion and confusion", *Topoi* 34 (1): 55–61. DOI: `10.1007/s11245-013-9175-x`

Priest, G., 2019, "Some comments and replies", pages 575–675 in C. Başkent and T. M. Ferguson (eds.), *Graham Priest on Dialetheism and Paraconsistency*, Springer: Berlin. DOI: `10.1007/978-3-030-25365-3_27`

Ripley, D., 2013, "Paradoxes and failures of cut", *Australasian Journal of Philosophy* 91 (1): 139–164. DOI: `10.1080/00048402.2011.630010`

Ripley, D., 2015a, "Anything goes", *Topoi* 34 (1): 25–36. DOI: `10.1007/s11245-014-9261-8`

Ripley, D., 2015b, "Comparing substructural theories of truth", *Ergo* 2: 299–328. DOI: `10.3998/ergo.12405314.0002.013`

Shapiro, L., 2011, "Deflating logical consequence", *The Philosophical Quarterly* 61: 320–342. DOI: `10.1111/j.1467-9213.2010.678.x`

Shapiro, L., 2013, "Validity Curry strengthened", *Thought* 2: 100–107. DOI: `10.1002/tht3.80`

Shapiro, S., 2002, "Incompleteness and inconsistency", *Mind* 111: 817–832. DOI: `10.1093/mind/111.444.817`

Shapiro, S., 2019, "Inconsistency and incompleteness, revisited", pages 469–479 in C. Başkent and T. M. Ferguson (eds.), *Graham Priest on Dialetheism and Paraconsistency*, Springer: Berlin. DOI: `10.1007/978-3-030-25365-3_22`

Smiley, T., 1957, "Entailment and deducibility", *Proceedings of the Aristotelian Society* 59: 233–254.

Tanswell, F. S., 2016, "Saving proof from paradox: Gödel's paradox and the inconsistency of informal mathematics", pages 159–173 in H. Andreas and P. Verdée (eds.), *Logical Studies of Paraconsistent Reasoning in Science and Mathematics*, Springer: Berlin. DOI: `10.1007/978-3-319-40220-8_11`

Weber, Z., 2014, "Naïve validity", *Philosophical Quarterly* 64 (254): 99–114. DOI: `10.1093/pq/pqt016`

Weir, A., 2005, "Naïve truth and sophisticated logic", pages 218–249 in J. Beall and B. Armour-Garb (eds.), *Deflationism and Paradox*, Oxford University Press: Oxford.

Whittle, B., 2004, "Dialetheism, logical consequence and hierarchy", *Analysis* 64 (4): 318–326. DOI: `10.1093/analys/64.4.318`

Zardini, E., 2011, "Truth without contra(di)ction", *Review of Symbolic Logic* 4: 498–535. DOI: `10.1017/S1755020311000177`

Zardini, E., 2012, "Näive modus ponens", *Journal of Philosophical Logic* 42 (4): 575–593. DOI: `10.1007/s10992-012-9239-1`

Zardini, E., 2013a, "It is not the case that [*p* and 'it is not the case that *p*' is true] nor is it the case that [*p* and '*p*' is not true]", *Thought* 1 (4): 309–319.

Zardini, E., 2013b, "Näive logical properties and structural properties", *The Journal of Philosophy* 110 (11): 633–644. DOI: `10.5840/jphil2013110118`

Zardini, E., 2015, "Getting one for two, or the contractors' bad deal. Towards a unified solution to the semantic paradoxes", pages 461-493 in T. Achourioti, H. Galinon, J. Martinez-Fernandez and F. Fujimoto (eds.), *Unifying the Philosophy of Truth*, Springer: Berlin. DOI: `10.1007/978-94-017-9673-6_23`

Massimiliano Carrara and Enrico Martino
Department of Philosophy, Sociology, Education and Applied
Psychology (FISPPA); University of Padua
Piazza Capitaniato 3, 35139 Padova, Italy
`{massimiliano.carrara, martinoenri}@gmail.com`