논리연구 20-2(2017) pp. 241-271

Can Gödel's Incompleteness Theorem be a Ground for Dialetheism?*

Seungrak Choi

【Abstract】 Dialetheism is the view that there exists a true contradiction. This paper ventures to suggest that Priest's argument for Dialetheism from Gödel's theorem is unconvincing as the lesson of Gödel's proof (or Rosser's proof) is that any sufficiently strong theories of arithmetic cannot be both complete and consistent. In addition, a contradiction is derivable in Priest's inconsistent and complete arithmetic. An alternative argument for Dialetheism is given by applying Gödel sentence to the inconsistent and complete theory of arithmetic. We argue, however, that the alternative argument raises a circularity problem. In sum, Gödel's and its related theorem merely show the relation between a complete and a consistent theory. A contradiction derived by the application of Gödel sentence has the value of true sentences, i.e. the both-value, only under the inconsistent models for arithmetic. Without having the assumption of inconsistency or completeness, a true contradiction is not derivable from the application of Gödel sentence. Hence, Gödel's and its related theorem never can be a ground for Dialetheism.

【Key Words】 Gödel's incompleteness theorem, Rosser's incompleteness theorem, Dialetheism, Inconsistent Arithmetic.

Received: Jan. 9, 2017 Revised: Jun. 2, 2017 Accepted: Jun. 5, 2017.
* I am grateful to the anonymous referees for comments on the first draft of the paper submitted, which helped a great deal of improvement.

1. Introduction

A conventional conception of inconsistency in a classical arithmetic tells that inconsistency implies everything. *Ex Contradictione Quodlibet*(hereafter *ECQ*), which means a contradictory premise leads any (true) sentences, often describes the classical inconsistency. Paraconsistent logicians, including Robert Meyer, Chris Mortensen, and Graham Priest, have challenged the orthodox view of the inconsistency with the arithmetical (inconsistent) structures that *ECQ* does not work as valid. Their counter-example to *ECQ* distinguishes between inconsistency and triviality, then gives an inconsistent but non-trivial structure for arithmetic. A theory Γ is inconsistent if Γ contains both a sentence φ and its negation $\sim\varphi$, otherwise Γ is consistent. Γ is trivial if Γ derives every (well-formed) sentence in the language L of Γ, otherwise Γ is non-trivial. The inconsistent and non-trivial arithmetic contains an instance of the form φ and $\sim\varphi$ (i.e. $\varphi \wedge \sim\varphi$), nonetheless $\varphi \wedge \sim\varphi$ does not imply every true sentence. If *ECQ* expresses that any instances of the form $\varphi \wedge \sim\varphi$ implies every true sentence, it has a counter-example and becomes an invalid rule as it does not preserve the truth-value of the premises to the conclusion in all arithmetical structures.

Meyer (1976) seems to have first suggested an inconsistent arithmetic and Meyer and Mortensen (1984) has further developed the inconsistent models of it. Priest (1997, 2000) has shown that all inconsistent models for arithmetic have a certain general form.[1)]

[1)] More precisely, there are two ways of preserving non-triviality of the inconsistent theory. Having *RM3*-models for arithmetic suggested in Meyer and Mortensen (1984), *RM3* assigns the values for the implication in the three-

Although they have a different counter-example to *ECQ*, the main reason for rejecting *ECQ* is of the same kind. Roughly put, their inconsistent arithmetic has a three-valued system taking the set of values – *true*, *both*, and *false*. If a sentence φ has a both-value, its negation $\sim\varphi$ either has the both-value. Henceforth, $\varphi \land \sim\varphi$ and $\sim(\varphi \land \sim\varphi)$ have the both-value. Inconsistent models for arithmetic regard true and both-value as the value of true sentences. As all sentences are not true and both in (some) inconsistent models, the truth value of $\varphi \land \sim\varphi$ does not preserve its true and both-value to all sentences. In other words, there exists a case that $\varphi \land \sim\varphi$ is a true sentence which has the both-value but its consequence is false.

A defender of *ECQ* may ask why we should accept the both-value in logic. Priest (2006a) has given the answer from paradoxes. He has strongly maintained Dialetheism which is the thesis that there exists a true contradiction. For convenience, we will call any sentences of the form $\varphi \land \sim\varphi$, 'contradiction', and any true contradiction in all discourses (or mathematical structures), 'dialetheia.' 'Dialetheism' intends to mean in this paper that there exists a dialetheia, in so far as no misapprehension appears. Priest keeps the view that any formalization of natural language semantics needs to take paradox into

---

valued Sugihara matrix {+1, 0, -1} which appears in Anderson and Belnap (1975). Let Φ, Ψ be a formula. Let $\neg$, $\land$, $\lor$, and $\rightarrow$ be a paraconsistent negation, conjunction, disjunction, and implication constant. Regarding *ECQ* with the form $(\Phi \land \neg\Phi) \rightarrow \Psi$, $(\Phi \land \neg\Phi) \rightarrow \Psi$ is not true in all valuations of *RM3*, so *ECQ* is not valid in *RM3*-models. On the other hand, Priest's inconsistent models in Priest(1979), called the Logic of Paradox (*LP*), define $\Phi \rightarrow \Psi$ as $\neg\Phi \lor \Psi$ and $\neg(\Phi \land \neg\Phi) \lor \Psi$ is true in all valuations of *LP*. However, *Modus Ponens* is not valid in *LP*, any sets of sentences true in *LP* are not trivial.

account. ‘This sentence is not true’ is a well-known sentence of the Liar Paradox. If it were true, then it is not true. Also, it is true if it were not true. Therefore, under the assumption that the suggested reasoning is sound, the liar sentence seems to be both true and not true. Any liar-type sentences are easily constructed in natural language. Priest has claimed that the paradoxical reasoning appeals us that there exists a dialetheia which has the both-value.

The main question is, ‘is there any dialetheia in arithmetic?’ Though natural language has any liar-type sentences, it is unclear that there exists a dialetheia in arithmetic. Priest (1979, 1984) and Priest (2006a: Ch.3 and Ch. 17) has proposed an argument for Dialetheism from Gödel’s first incompleteness theorem. Gödel’s theorem shows, for any consistent and sufficiently strong formal axiomatic theory Γ for arithmetic, there exists a Gödel sentence φ constructed in the language L of Γ but not provable in Γ. There may be a parallel between the Gödel sentence, which says ‘I am not provable’, and the Liar sentence. Were the Gödel sentence regarded as an arithmetical analogy of the Liar sentence, Gödel’s first incompleteness theorem is the consistent counterpart of the Liar Paradox. As Priest has claimed, if the Liar sentence could give a dialetheia, either could the Gödel sentence in Γ having no assumption of consistency of Γ. It seems to be that Priest presumes the inconsistency of our linguistic practice and attempts to derive the inconsistency of our naive proof procedures in natural language.[2] However, exactly how one is supposed to derive the

---

[2] I thank an anonymous reviewer for pointing out that the inconsistency of our linguistic practice in natural language can imply the inconsistency of our naive notion of proof.

inconsistency of our naive proof procedures from our linguistic practice remains unclear. Priest has found an answer from Gödel's theorem. The present paper investigates Priest's view on Gödel's theorem and his argument for Dialetheism given by applying Gödel sentence to our naive notion of proof.

The issue has been discussed in Charles Chihara (1984) and Neil Tennant (2004). Priest maintains that any correct formalization of our naive proof procedure is inconsistent and it is tantamount to show that a dialetheia exists. Hence, the main tension between them is the nature of naive proof procedures. For instance, Chihara (1984) examines Priest's view on Gödel's theorem with the following argument:

(1) If Γ has a complete formalization of our naive proof procedures, then only truths are provable in Γ.
(2) Γ has the complete formalization of our naive proof procedures.
(3) Therefore, by (1) and (2), only truths are provable in Γ.

Provided that Γ derives a contradiction, there exists a dialetheia since only truths are provable in Γ. Many logicians have maintained that not all mathematical proof procedures can be completely formalized. Chihara (1984) denies the assumption (2). Likewise, considering that there exists no dialetheia in intuitionistic logic, Tennant (2004) claims Gödel's theorem merely shows that we cannot have the complete characterization of our naive proof procedures. Their interpretation of Gödel's theorem, however, may be the consistent counterpart with the law of non-contradiction as Priest (1984, 2006a) denies. They may have a different conception of our naive proof procedures. The tension of their dispute may converge on the problem of whether the rejection of the law of non-contradiction is legitimate or not. In this paper, we set aside the issue of the nature of our naive notion of proof. Rather, we focus on Gödel's and its related proofs *per se*. The aim of the paper is not to devastate Dialetheism in arithmetical discourses but to claim that if we drop the assumption of consistency and of inconsistency, Gödel's theorem could not be a ground for Dialetheism.

In Section 2, we will argue that for a Gödel sentence φ, $\varphi \wedge \sim\varphi$ is derivable only in a complete theory of arithmetic. It shall be claimed that the lesson of Gödel's proof is that any sufficiently strong and intuitively correct arithmetic cannot be both complete and consistent. In other words, although $\varphi \wedge \sim\varphi$ is provable in a complete theory, Gödel's and its related theorem do not provide any clue that a theory of arithmetic must be complete. An expected answer from Priest is that there exists a decidable complete inconsistent arithmetic and so the complete arithmetic is provable. The argument for Dialetheism is achieved by the application of Gödel sentence to the inconsistent and complete theory of arithmetic. After introducing Priest's inconsistent models for arithmetic, in Section 3, we will argue that the circularity problem is waiting for the argument. In conclusion, we will query whether the inconsistent (or paraconsistent) mathematics needs Gödel's theorem as its motivation in that inconsistent logic and mathematics are achieved without Gödel's theorem.

2. Deriving a Contradiction from Gödel's Original Proof.

As we have noted, one of the Priest's main motivations for Dialetheism is Gödel's theorem. He applies Gödel sentence to a naive notion of proof in natural language and attempts to make an argument for Dialetheism. His 'naive proof' stands for the informal deductive arguments from basic sentences which are put forward to be maintained as true without any proof of it. Each axiom of Peano Arithmetic(hereafter *PA*) is to be a basic sentence. In this regard, his naive proof intends to mean any informal mathematical proof procedures from the axioms. Affirming his claim that a set of the naive proofs satisfies the conditions of Gödel's theorem and let Γ be a formalization of the naive proof procedures for arithmetic, he claims,

> … if [Γ] is consistent there is a sentence φ which is not provable in [Γ], but which we can establish as true by a naive proof, and hence *is* provable in [Γ]. The only way out of the problem, other than to accept the contradiction, and thus [D]ialetheism anyway, is to accept the inconsistency of naive proof. So we are forced to admit that our naive proof procedures are inconsistent. But our naive proof procedures just are those methods of deductive argument by which things are established as true. It follows that some contradictions are true; that is, [D]ialetheism is correct. (2006a: 44)

We now venture to suggest that his interpretation of Gödel's theorem is half true, as the moral of Gödel's proof is that it is unable to be proved by Γ itself that Γ is both complete and consistent.

To begin our story with the Liar Paradox which is one of the main motivations of Dialetheism. Let φ be a liar-type sentence, saying 'φ' is not true. A classical naive proof process entails an equivalent relation

$\varphi \equiv \sim\varphi$ such that φ is true if and only if $\sim\varphi$ is true and thus $\varphi \wedge \sim\varphi$ is true. If the consistency of Γ is a primary criterion for a legitimate theory Γ, Γ would exclude a liar-type sentence, φ, which leads a contradiction, $\varphi \wedge \sim\varphi$. On the other hand, Γ could not single out φ, if Γ is complete with its language L having φ. The definition of 'complete' and 'consistent' runs as follows. In accordance with standard practice, we write '$\Gamma \vdash \varphi$' to mean that Γ derives φ and '$\Gamma \nvdash \varphi$' means that Γ does not derive φ.

Definition 1. *Let Γ be any theory and L be a language. (1) Γ is complete if for each sentence φ in L, either $\Gamma \vdash \varphi$ or $\Gamma \vdash \sim\varphi$.*[3] *(2) Γ is consistent if there exists no φ such that $\Gamma \vdash \varphi$ and $\Gamma \vdash \sim\varphi$.*

The liar-type sentence φ in L implies $\sim\varphi$ and vice versa. (i.e. $\varphi \equiv \sim\varphi$.) If Γ is complete, $\Gamma \vdash \varphi \wedge \sim\varphi$. Hence, Γ is not consistent. On the other hand, if Γ is consistent, then there is no φ such that $\Gamma \vdash \varphi$ and $\Gamma \vdash \sim\varphi$. Since $\varphi \equiv \sim\varphi$, the assumption that $\Gamma \vdash \varphi$ leads $\Gamma \vdash \sim\varphi$. This contradicts the consistency of Γ, so we have $\Gamma \nvdash \varphi$. The similar process provides $\Gamma \nvdash \sim\varphi$. Therefore, Γ is not complete. The assumption that Γ is consistent can be rejected if $\Gamma \vdash \varphi \wedge \sim\varphi$. For there is no clue that Γ is complete, however, neither $\Gamma \nvdash \varphi$ implies $\Gamma \vdash \sim\varphi$ nor $\Gamma \nvdash \sim\varphi$ implies $\Gamma \vdash \varphi$. With the language L having φ, it seems, Γ cannot be both complete and consistent. With regard to Gödel's theorem, we will arrive at the same conclusion.

---

[3] We will use 'completeness' in this sense if there be no misapprehension.

Priest (2006a, p.5) says that consistency of our linguistic principles cannot be invoked as a regulatory one against inconsistency. Rather, he assumes that inconsistency is the natural presupposition of the principles. While putting aside the presupposition of consistency, a complete theory Γ derives a contradiction. However, without supposing inconsistency or completeness, there is no reason that Γ has a dialetheia. The similar result of which we mentioned above can be given by Gödel's theorem.

Gödel's first incompleteness theorem shows that, for any consistent and sufficiently strong theory Γ of arithmetic, there exists a sentence φ in the language L of Γ but not derivable in Γ. The result would not be shown in all theories of arithmetic. It may be proved only by the theory that can represent all primitive recursive functions. Roughly, to say that a function is recursive is to say that there exists an effective calculable method to decide its value. A function is primitive recursive if it can be obtained from the basic functions, such as zero, successor and the various identity functions, by composition and recursion. Many arithmetical truths can be formulated by the primitive recursive functions. When a theory has an ability to express certain truths through all primitive recursive functions, we say that it can represent all primitive recursive functions. If every (naive) proof procedure can be characterized by all primitive recursive functions in the theory, all those are representable. One of the interpretations of Gödel's theorem might be that, for any given theory Γ for arithmetic which all primitive recursive functions are representable, Γ is unable to express all arithmetical truths. Whereas Priest claims that it merely shows an inconsistency of Γ. His proof in Priest (2006a, pp. 48-60) starts from the consistency assumption of Γ and derives, for a Gödel sentence φ, that it is not the case that φ and φ is true. However, from his proof, φ and $\sim\varphi$ is not to be derivable from Γ unless the proof presumes the completeness of Γ. In Priest's perspective on the linguistic principles, it may be begging the question to assume that Γ is consistent. If it is, either we should drop the inconsistency assumption of Γ. Therefore, it is desirable to set aside both the presuppositions of consistency and of inconsistency, and to show whether Γ is inconsistent or incomplete. Without those presuppositions, we only have a relation between the complete and the consistent theory from Gödel's proof. With some terminologies, lemma, and theorem, we give a proof of a contradiction true in ω-complete and complete theory Γ.

Let φ be any formula, $\lceil\varphi\rceil$ be the code of φ and, for any given natural number *n*, $\overline{n}$ be its numeral. Hence, $\overline{\lceil\varphi\rceil}$ is the numeral of the code of φ. Instead of assuming (ω-)consistency of Γ, the assumption of the ω-completeness and the completeness of Γ can be taken for our proof of $\Gamma \vdash \varphi \wedge \sim\varphi$.[4] Having the definitions of 'ω-complete' and 'ω-consistent', we can derive a contradiction true in the ω-complete and complete theory of arithmetic. A sketch of the following result can be found in Appendix A.

Definition 2. *Let L be a language of arithmetic and Γ be any theory in L. (1) Γ is ω-complete if $\Gamma \vdash \exists x\varphi(x)$ implies $\Gamma \vdash \varphi(\overline{n})$ for some natural number n. (2) Γ is ω-consistent if there exists no φ such that $\Gamma \vdash \exists x\varphi(x)$ and $\Gamma \vdash \sim\varphi(\overline{n})$ for all n.*

[4] As Tarski(1933) has investigated, ω-consistency implies consistency. We drop the ω-consistency assumption.

Theorem 1. *Let Γ be any given ω-complete theory of arithmetic that can represent all primitive recursive functions. If Γ is complete, there exists φ such that*

$$\Gamma \vdash \varphi \wedge \sim\varphi.$$

Proof. See Appendix A.

Theorem 2. (Gödel, 1931) *If Γ is ω-consistent, Γ is not complete.*

With the assumption of ω-consistency, Gödel (1931) shows the first incompleteness theorem, i.e. Theorem 2. Theorem 1 and 2 show that in the language, *L,* having Gödel sentence φ, Γ cannot be both complete and (ω-)consistent. $\varphi \wedge \sim\varphi$ is not derivable from Γ unless it is assumed that Γ is complete. Without the completeness assumption of Γ, $\Gamma \nvdash \varphi$ (or $\sim\varphi$) does not mean $\Gamma \vdash \sim\varphi$ (or φ). Not all (inconsistent and non-trivial) theories of arithmetic are complete. (Cf. Meyer and Mortensen, 1984.) Gödel's proof and its application merely give us the lesson that only in the complete theory Γ there exists a sentence $\varphi \wedge \sim\varphi$ true in Γ.[5)] In other words, Gödel's first incompleteness theorem (Theorem 2) does not say that Γ is inconsistent. Even though we consider Theorem 1, we cannot claim

[5)] There is another proof of the incompleteness theorem. Barkley Rosser (1936) shows that the assumption of ω-consistency in Gödel's first incompleteness theorem can be replaced by consistency. Accepting Rosser's incompleteness proof, one may argue that Γ has a dialetheia at the expense of consistency. Although we follow the line of Rosser's proof, we should assume the completeness of Γ to derive a contradiction. The precise proof of it is in the Appendix B.

that a *general theory* of arithmetic is inconsistent, as all theories of arithmetic are not complete.

An expected answer from this objection is that Priest purports to have proved an inconsistency in Γ, which cannot be dismissed by asserting that Γ is consistent. In addition, he may claim that the completeness of an inconsistent arithmetic is provable and the inconsistent arithmetic is more general than a classical arithmetic. Truly, there is a decidable complete inconsistent arithmetic. However, if a dialetheia were a contradiction true in all mathematical structures, Gödel's theorem (Theorem 1) as the main motivation for Dialetheism should support that any theories (or any *general theory*) of arithmetic have a contradiction. For only in some complete theories of arithmetic, a contradiction is derivable, it has to be shown that any theories of arithmetic are complete and have a dialetheia. Though we restrict our scope of theories into *some* theories, Theorem 1 and 2 does not ensure the completeness of the theories. It is unconvincing that Gödel's and its related results are the ground for Dialetheism.

The completeness of (some) inconsistent arithmetic can be proved by constructing an inconsistent model, $\mathcal{N}$, which extends any consistent models, $\mathcal{M}$, of classical arithmetic. Considering that the set $Th(\mathcal{N})$ of sentences true in $\mathcal{N}$ is complete. The application of the Theorem 1 to $Th(\mathcal{N})$ yields a contradiction. Priest(2006a, pp. 236-237), in practice, makes an argument for Dialetheism or for the inconsistent arithmetic in this way. The application of the Gödel sentence to $Th(\mathcal{N})$ seems to show that Dialetheism is true. The simple argument for Dialetheism is that, for a Gödel sentence φ, if $Th(\mathcal{N})$ is complete, then $Th(\mathcal{N}) \vdash \varphi \wedge \sim\varphi$ and so Dialetheism is true. The problem is that to make $Th(\mathcal{N})$ complete, there must be an inconsistent object in a domain of $\mathcal{N}$ and its interpretation should have a both-value for dialetheias. To have the inconsistent objects(or the both-value), Dialetheism is to be true. Unfortunately, because, for Priest, Gödel's theorem supports Dialetheism, his argument falls into a circularity problem. In the next section, we shall investigate his inconsistent arithmetic and argue that his argument involves the problem of circularity.

3. Priest's Inconsistent Models for Arithmetic and the Circularity Problem.

In this section, we shall investigate Priest's inconsistent and complete arithmetic. Not all the inconsistent theories for arithmetic are complete. Priest's inconsistent models for arithmetic are produced by the Collapsing Theorem which implies that a set of sentences true in a collapsed model is complete. To this end, we firstly introduce Priest's inconsistent models and the Collapsing Theorem. An inconsistent and complete theory of arithmetic is a direct consequence of the Collapsing Theorem. Next, we will carry on our discussion of Priest's argument for Dialetheism from Gödel's theorem (Theorem 1) and argue that his argument raises a circularity problem.

Priest (1997, 2000) has proposed inconsistent models for arithmetic setting out in the Logic of Paradox, *LP*. *LP* interpretation suggested in Priest (1979, 1991) is based on three values; *true*, *both* and *false*. The language *L* of *LP* is that of classical first-order logic, including function symbols and identity. The *LP* interpretation (or structure) $\mathcal{N}$ for *L* is a pair $< D, I >$, where *D* is a non-empty set and *I* assigns denotations to the non-logical symbols of *L* in the following way.

- For any constant symbol *d*, *I(d)* is a member of *D*.
- For every *n*-ary function symbol *f*, *I(f)* is an *n*-ary function on *D*.
- For every *n*-ary predicate symbol φ, $I(\varphi)$ is the pair $< I^{+}(\varphi), I^{-}(\varphi) >$ where $I^{+}(\varphi)$ and $I^{-}(\varphi)$ are the extension and anti-extension of φ respectively.

We should note that, for any *n*-ary predicate φ, $I^{+}(\varphi) \cup I^{-}(\varphi) = \{< d_1, \ldots, d_n >; d_1, \ldots, d_n \in D\}$ but no need to be $I^{+}(\varphi) \cap I^{-}(\varphi) = \emptyset$. If the extension and anti-extension of a predicate are disjoint in the *LP* interpretation, we shall call it classical structure, $\mathcal{M}$.

Finally, let $\mathscr{v}$ be a valuation from the formulas to truth values where $\mathscr{v}(\Phi) \in \{\{1\}, \{1,0\}, \{0\}\}$. Where φ is a predicate and $t_1, \ldots, t_n$ are terms, the valuations for atomic formulas, negation($\neg$), conjunction($\wedge$), and the universal quantifier are as follows[6]:

- For an n-ary predicate φ,
  $1 \in \mathscr{v}(\varphi(t_1, \ldots, t_n))$ iff $< I(t_1), \ldots, I(t_n) > \in I^{+}(\varphi)$,
  $0 \in \mathscr{v}(\varphi(t_1, \ldots, t_n))$ iff $< I(t_1), \ldots, I(t_n) > \in I^{-}(\varphi)$.
- For a formula Φ, Ψ of *L*,
  $1 \in \mathscr{v}(\neg\Phi)$ iff $0 \in \mathscr{v}(\Phi)$,

6) Paraconsistent logic has a different use of the implication and negation from that of classical logic. We will use '⊃', '∼', '≡' for the material implication, classical negation, and classical equivalence relation respectively. Also, '→', '¬', '↔' will be used, respectively, for the paraconsistent implication, negation, and equivalence relation. 'iff' is the abbreviation of 'if and only if'.

$0 \in v(\neg\Phi)$ iff $1 \in v(\Phi)$,
$1 \in v(\Phi \wedge \Psi)$ iff $1 \in v(\Phi)$ and $1 \in v(\Psi)$,
$0 \in v(\Phi \wedge \Psi)$ iff $0 \in v(\Phi)$ or $0 \in v(\Psi)$,
$1 \in v(\forall x\Phi)$ iff $1 \in v(\Phi[x/d])$ for all $d \in D$ where $[x/d]$ means the substitution of *d* for *x* in Φ,
$0 \in v(\forall x\Phi)$ iff $0 \in v(\Phi[x/d])$ for some $d \in D$.

The valuations for disjunction and quantification can be taken as defined by the suggested valuation above. That is to say, $v(\Phi \vee \Psi) = v(\neg(\neg\Phi \wedge \neg\Psi))$, $v(\exists x\Phi) = v(\neg\forall x\neg\Phi)$, and, as usual, $v(\Phi \rightarrow \Psi) = v(\neg\Phi \vee \Psi)$. Truth conditions for classical logic are obtained by ignoring the second clause of each connective. The above interpretation extends to equality in the following sense.

Definition 3. *For any given LP-interpretation I and J, J is an extension of I iff for every predicate* φ, $I^+(\varphi) \subseteq J^+(\varphi)$ *and* $I^-(\varphi) \subseteq J^-(\varphi)$.
Theorem 3. *Let I, J be (LP-)interpretations and J is an extension of I. Let* v_1, v_2 *are valuations for I, J respectively. For any formula* $\Phi, v_1(\Phi) \subseteq v_2(\Phi)$.
Proof. See Priest(1997).

Priest(1997, 2000) takes the names to be the members of *D* themselves and adopts the convention that for every $d \in D, I(d)$ is just *d* itself. Let $L_{\mathcal{N}}$ be a language of $\{0, S, +, \cdot\}$ for (inconsistent) arithmetic augmented from *L* with a name for every member of *D* of $\mathcal{N}$ and Γ be a theory in $L_{\mathcal{N}}$ extending *PA*. Suppose $\mathcal{N}$ be a non-standard model of Γ and $\approx$ be a congruence relation on $\mathcal{N}$ with only finitely many equivalence classes.[7] We define $D^{\approx}$ to be the set of equivalence classes and say $[d]$ is the equivalence class of d in D under $\approx$. So to speak, $[d]$ is defined as $\{x; x \approx d\}$ and $D^{\approx} = \{[d]; d \in D\}$. A new interpretation, $\mathcal{N}^{\approx} =< D^{\approx}, I^{\approx} >$, called the *collapsed interpretation* is given as follows:

- For every constant d, $I^{\approx}(d) = [I(d)]$.
- For every n-place function f,
$$I^{\approx}(f)([d_1], \dots, [d_n]) = [I(f)(d_1, \dots, d_n)].$$[8]
- For every predicate φ and $1 \leq i \leq n$,
$< [d_1], \dots, [d_n] > \in I^{\approx}_{+}(\varphi)$ iff for some $e_i \approx d_i$, $< [e_1], \dots, [e_n] > \in I^{+}(\varphi)$,
$< [d_1], \dots, [d_n] > \in I^{\approx}_{-}(\varphi)$ iff for some $e_i \approx d_i$, $< [e_1], \dots, [e_n] > \in I^{-}(\varphi)$.
- $I^{\approx}_{+}([x] = [y]) = \{< [x], [y] >; x \approx y\}$,
- $I^{\approx}_{-}([x] = [y]) = \{< [x], [y] >; x \neq y\}$.

It is easily seen that $I^{\approx}$ is an extension of I. Having the collapsed interpretation $\mathcal{N}^{\approx}$, we have the Collapsing theorem.

---

7) $\approx$ is also an equivalence relation which satisfies that if $d_i \approx e_i$ for all $1 \leq i \leq n$, then $I(f)(d_1, \dots, d_n) \approx I(f)(e_1, \dots, e_n)$ where f is an n-place function in $L_{\mathcal{N}}$ and $d_i, e_i \in D$. We only focus on the finite *LP* models for Γ in order to set aside the problems of the infinite *LP*. Priest(1997) claims the existence of a further family of finite *LP* models of Γ, called 'clique models', with his intended congruence relation. Paris and Pathmanathan (2006) indicates that Priest's proof, stating his intended congruence relation, has an error, and Paris and Sirokfskich (2008) extend their work into the infinite. Regardless of the cases of the infinite *LP* models, we can discuss the completeness and inconsistency of Γ. We leave this issue aside in the present paper.

8) In the same way, $I^{\approx}([x] + [y]) = [I(x + y)]$, $I^{\approx}([x] \cdot [y]) = [I(x \cdot y)]$, $I^{\approx}(S([x])) = [I(S(x))]$.

Theorem 4. (the Collapsing Theorem) *Let* v *be a valuation for* $\mathcal{N}$ *and* $v^{\approx}$ *for* $\mathcal{N}^{\approx}$*. For any formula* Φ *of* $L_{\mathcal{N}}$*,* $v(\Phi) \subseteq v^{\approx}(\Phi)$*.*
Proof. Priest(1997).

In virtue of the collapsed interpretation, a formula Φ is (*LP*-)logical truth iff every interpretation is a model for it. The Collapse Theorem tells us that every *LP*-logical truth is a logical truth of classical first-order logic, but not vice versa. Let $Th(\mathcal{N}^{\approx})$ be a set of sentences true in $\mathcal{N}^{\approx}$. Priest(1994) regards that the completeness and inconsistency of $Th(\mathcal{N}^{\approx})$ are immediate consequences of the Collapsing Theorem.

With regard to the inconsistency of $Th(\mathcal{N}^{\approx})$, $\exists x(x = x \wedge x \neq x)$ holds in a collapsed model of $Th(\mathcal{N}^{\approx})$. Let $\mathcal{M} =< D', I' >$ be any classical model of $Th(\mathcal{N}^{\approx})$ and $\approx$ be an equivalence relation on D' which is either a congruence relation in terms of the interpretations of the function symbols. Accepting the collapsed interpretation above, we produce a collapsed interpretation $\mathcal{M}^{\approx}$. Some classically distinct numbers in $\mathcal{M}$ are collapsed into equivalence classes and the equivalence classes preserve the non-identities of their members. In other words, if x and y are distinct numbers, and in the equivalence class $[z]$, $[x] = [y]$ but $[x] \neq [y]$ since $x \approx y$ but $x \neq y$. Hence, $\exists x(x = x \wedge x \neq x)$ holds in $\mathcal{M}^{\approx}$.[9)]

The next thing is about the completeness of $Th(\mathcal{N}^{\approx})$. We assume the ω-completeness of $Th(\mathcal{N}^{\approx})$ for the sake of convenience. In the finite *LP*-models of arithmetic, even numerical equations can be

---

9) *LP*-collapse models are quotient algebras of classical arithmetic which produce diverse models similar to modular arithmetic. The reader can consult Priest (1997) for the detailed explanations and the examples of inconsistent models for arithmetic, such as *cyclic* and *heap* models.

inconsistent. For the truth-values of a sentence ψ and $\neg\psi$ are determined by different procedures I^+ and I^-. In this sense, both the Gödel sentence, φ, and its negation, $\neg\varphi$, have the truth-value. Let define a provability predicate for the Gödel sentence φ in $Th(\mathcal{N}^{\approx})$, $Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)$, as $\exists y Prf_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner, y)$ which says that φ is provable in $Th(\mathcal{N}^{\approx})$. Provided that $L_{\mathcal{N}^{\approx}}$ has φ and $Th(\mathcal{N}^{\approx}) \vdash \varphi$, there is an n such that $Prf_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner, n)$ is provable in $Th(\mathcal{N}^{\approx})$. Since $Prf_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner, n)$ is true in $\mathcal{N}^{\approx}$, $I^+\big(Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)\big) = n$. By the Collapsing Theorem, (3) in Appendix A is true in $\mathcal{N}^{\approx}$ and $\varphi \leftrightarrow \neg Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)$ holds in $\mathcal{N}^{\approx}$.[10] It follows, by the assumption $Th(\mathcal{N}^{\approx}) \vdash \varphi$, that $Th(\mathcal{N}^{\approx}) \vdash \neg\overline{Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)}$. If $\neg Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)$ is true in $\mathcal{N}^{\approx}$ and only we have an interpretation I^+ with its extension, we should say that there exists an n such that $I^+\big(Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)\big) = n$ and such n does not exist. Taking the similar example, Stewart Shapiro (2002) has refuted Priest's Dialetheism. Shapiro(2002, p. 828) asks '[h]ow can the dialetheist go on to maintain that … [n] is *not* the code of a [$Th(\mathcal{N}^{\approx})$]-derivation of [$\varphi$]?' However, he lacks the point

---

[10] Let Φ, Ψ be formulas and $\circ$ be constant. We call $\circ$ 'detachable constant' if $\Phi \circ \Psi$ and Φ jointly imply Ψ. Say that the detachment inference is any form of inferences that if $\Phi \circ \Psi$ and Φ then Ψ. Beall and Foster and Seligman (2012) argues that *LP* does not admit the detachment inference. Without the detachment, it seems that the equivalence relation between the Gödel sentence, φ, and its negation, $\neg\varphi$ does not hold. Thus, in *LP*, one can reject the equivalence relation $\varphi \leftrightarrow \neg\varphi$. However, Priest (1991) and Priest (2006, Ch.8 and Ch. 16) accept the detachment inference as a quasi-valid inference which means it is classically valid but dialetheically invalid in a minimally inconsistent *LP*. Also if φ is true in a classical model, then, in the collapsed model $\mathcal{N}^{\approx}$, φ is true, by the Collapsing Theorem. To avoid an unnecessary dispute, we assume in this section that $\varphi \leftrightarrow \neg\exists y Prf_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner, y)$ and $\varphi \leftrightarrow \neg\varphi$ are true.

that Priest's inconsistent models have two interpretation I^{+} and I^{-}. As Priest (2006a, p. 242) puts, $\mathcal{N}^{\approx}$ has an additional interpretation I^{-} having the anti-extension and so $I^{-}\big(Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)\big) = m$ for some *m* in $\mathcal{N}^{\approx}$. $Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)$ and $\neg Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)$ are equivalent in ω-complete theory $Th(\mathcal{N}^{\approx})$.[11] It follows that $I^{\approx}\big(Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)\big) = I^{\approx}\big(\neg Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)\big)$. Let $[g]$ be the equivalence class containing *m* and *n*. As *m* and *n* are in $[g]$, $I_{+}^{\approx}\big(Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)\big) \cap I_{-}^{\approx}\big(Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)\big) = [g]$ and so $Th(\mathcal{N}^{\approx})$ is not classical. Unlike any consistent classical models for $Th(\mathcal{N}^{\approx})$, every name in $L_{\mathcal{N}}$ has its denotations and $\mathcal{N}^{\approx}$ has an inconsistent object $[g]$. Moreover, $\mathcal{N}^{\approx}$ has separate sets of extensions with $I_{+}^{\approx}$ and $I_{-}^{\approx}$. To determine whether φ is true in $\mathcal{N}^{\approx}$, we have to look to see whether $[g] \in I_{+}^{\approx}(\varphi)$. To determine whether $\neg\varphi$ is true in $\mathcal{N}^{\approx}$, we have to look to see whether $[g] \in I_{-}^{\approx}(\varphi)$. The separate processes of truth-value determination with the inconsistent object $[g]$ make $Th(\mathcal{N}^{\approx})$ complete because both the Gödel sentence, φ, and its negation, $\neg\varphi$, have the truth-value in $\mathcal{N}^{\approx}$ and $\varphi, \neg\varphi \in Th(\mathcal{N}^{\approx})$.

About the completeness of $Th(\mathcal{N}^{\approx})$, $Th(\mathcal{N}^{\approx})$ contains the Gödel sentence φ and $\neg\varphi$. Consider any classical models, $\mathcal{M} =< D', I' >$, for $Th(\mathcal{N}^{\approx})$, having no anti-extension, a set $Th(\mathcal{M})$ of sentences true in $\mathcal{M}$ neither have the Gödel sentence φ nor $\neg\varphi$. $Th(\mathcal{M})$ is incomplete but $Th(\mathcal{N}^{\approx})$ is complete. $\mathcal{N}^{\approx}$ is a collapsed model produced by applying the collapsed interpretation to any (consistent) classical model $\mathcal{M}$ for arithmetic. In other words, adding inconsistent

---

[11] Assuming $Th(\mathcal{N}^{\approx})$ is ω-complete. A simple variation of the claim 1 and 2 without 3 of Theorem 1, $Th(\mathcal{N}^{\approx}) \vdash \varphi \leftrightarrow \neg\varphi$. We already know that $\varphi \leftrightarrow \neg Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)$ is provable in $Th(\mathcal{N}^{\approx})$. Either $\varphi \leftrightarrow Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)$ is. Hence, $Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner) \leftrightarrow \neg Prv_{Th(\mathcal{N}^{\approx})}(\ulcorner\varphi\urcorner)$ is provable in $Th(\mathcal{N}^{\approx})$.

objects to $\mathcal{M}$ gives an inconsistent *LP* model for arithmetic and makes $Th(\mathcal{N}^{\approx})$ complete.

The last thing we should talk is the circularity problem of Priest's argument for Dialetheism from Gödel's theorem (Theorem 1). $Th(\mathcal{N}^{\approx})$ is complete by the collapsed interpretation. As Priest (2006a, pp. 236-237) has argued, an application of Theorem 1 to $Th(\mathcal{N}^{\approx})$ leads a contradiction, $\varphi \wedge \neg\varphi$. A derivation of the contradiction seems to guide us the view that Gödel's theorem may be a ground for Dialetheism. Although we merely accept Theorem 1 but not Theorem 2, his argument raises the circularity problem. The thing is that to make $Th(\mathcal{N}^{\approx})$ complete we should extend a (consistent) classical model $\mathcal{M}$ for *PA* with the extension for φ and the anti-extension for $\neg\varphi$. It means that we need a both-value for $\varphi \wedge \neg\varphi$ in order to make $Th(\mathcal{N}^{\approx})$ complete. As the interpretation I^{+} and I^{-} with inconsistent objects in $\mathcal{N}^{\approx}$ provides the both-value in $v^{\approx}$, φ and $\neg\varphi$ have the both-value, then $Th(\mathcal{N}^{\approx})$ becomes complete. To have the inconsistent objects in $\mathcal{N}^{\approx}$, however, Dialetheism must be true. Priest has given an argument for Dialetheism from Gödel's theorem, and so we consider that Theorem 1 supports Dialetheism. As noted in Section 2, Theorem 1 does not show the inconsistency of a certain theory Γ of arithmetic. It has a conditional form that if Γ is complete, then Γ is inconsistent. To support Dialetheism from Theorem 1, we have to prove that a theory of arithmetic, e.g. $Th(\mathcal{N}^{\approx})$, is complete. Circularity arises. Therefore, Priest's argument for Dialetheism from Gödel's theorem has a circularity problem.

Priest may answer that the existence of inconsistent objects or the both-value are guaranteed by the fact that our expressions in natural language and its linguistic principles are inconsistent, and so his argument is not circular. Unfortunately, such an answer is nothing but a rejection of the view that Gödel's theorem is a ground for Dialetheism.

The meaning of '=' in Priest's inconsistent arithmetic, $Th(\mathcal{N}^{\approx})$, is ambiguous. Both the congruence relation and the identity relation are regarded as the meaning of '='. As ambiguous expressions in natural language often lead an inconsistency, an ambiguous use of '=' derives a contradiction. However, Gödel's theorem does not give any reason for the legitimacy of the ambiguous use of '='. As we have discussed in Section 2, Theorem 1 and 2 neither show incompleteness nor inconsistency if we drop the assumption of consistency and of completeness of a theory. Gödel's and its related theorem only explain the relation between consistent and complete theories. If the inconsistency of natural language implies the existence of a dialetheia in arithmetic, Dialetheism is supported by the inconsistency of natural language, but not by Gödel's theorem (Theorem 1).

4. Conclusion: Do Paraconsistent Logicians Need Gödel's Theorem as Its Motivation?

There is more than one philosophical significance of Gödel's Theorem. Some classical logicians have believed that the implication of Gödel's Theorem is the inequality of the notion of *truth* in the standard interpretation and *provability* in a formal system. For them, truth transcends a (formal) proof. An intuitionist has denied the notion of the classical truth and claimed the truth as knowable(or provable).

Especially, when he has rejected the view that Gödel's theorem is not a ground for Dialetheism, Tennant(2004, p. 383) says, following Dummett(1963), '[w]hat Gödel's Theorem show is that we can never once-and-for-all delimit … the resources of 'naive provability''. These philosophical significances are the consistent counterpart of Gödel's theorem because classicists and intuitionists have never concerned about the inconsistent and non-trivial theories for arithmetic. There might be a paraconsistent significance of Gödel's Theorem. The complete counterpart of Gödel's theorem may show the existence of a true contradiction in a complete theory of arithmetic. We should not lose the point that only in the complete theory a contradiction is derivable as a theorem, on account of Gödel's and its related theorems. Dialetheism is not a promising paraconsistent significance of Gödel's Theorem if Dialetheism is the view that there exists a true contradiction in *all* (mathematical) structures. Furthermore, to avoid the circularity problem argued in Section 3, if Priest presumes an inconsistency of the linguistic principles of natural language, a derivation of the contradiction from Gödel's proof would rather be a consequence of the inconsistent and complete arithmetic than being a ground of it. In this sense, Gödel's theorem (Theorem 1) is not a ground for Dialetheism.

The other option is to accept mathematical pluralism which is the doctrine that there are different mathematical structures where distinct and incompatible theorems and laws hold. (Cf. Geoffrey Hellman and John Bell (2006) and Priest (2013)). Priest (2013) may consider a full-blooded platonist's version of mathematical pluralism to be the noneist position sketched in Priest (2005, Ch. 7). He points out that the inconsistent mathematics adds further to diverse the consistent mathematics. Given the perspective of Priest (2005), he accepts many worlds as well as impossible ones. Also, having the view of Priest (2013), logic may differ from world to world. Since logical truths vary across different worlds, he may accept alethic pluralism about truth which says that there exists more than one truth property. In some (impossible) worlds, a contradiction can be true and so Dialetheism is true, but this argument for Dialetheism is from the noneist's version of mathematical pluralism, but not from Gödel's Theorem.[12)] The derivation of the contradiction is merely a consequence of the inconsistent linguistic principles, but not a ground of Dialetheism.

Last but not the least, not all paraconsistent logicians accept Gödel's theorem as its motivation. Meyer and Mortensen (1983, p. 924) has shown the incompleteness of relevant arithmetic including his $R^{\#}$, $R^{\#\#}$, $RM^{\#}$, and $RM^{\#\#}$, without Gödel's incompleteness theorem. Moreover, Meyer (1996) rejects the formalization methods related with '~' in Gödel's proof as dirty tricks because '~' in Gödel's proof has a different meaning of 'not' in English. Without Gödel's and its related theorem, the paraconsistent logic and its relevant arithmetic are conceived. All in all, Gödel's theorem cannot be the ground for

---

12) Priest seems to take a double face in the discussion of pluralism. Priest(2006b, pp. 206-207) attacks on alethic pluralism. In this sense, his dialetheia seems to have only one property of truth that should be true in all *correct* theories of arithmetic (or all mathematical structures). In the stance of Priest (2006b), he can claim that there is only one actual world, only one actual truth and the actual world has an inconsistent linguistic principle of its natural language. If he regards that the actual world is inconsistent, the ground of the inconsistency or the existence of dialetheia is not from Gödel's Theorem, but from the inconsistency of the actual world.

Dialetheism, and, for some paraconsistent logicians, it does not need to be.

Appendix A. Deriving a Contradiction from Gödel's Proof.

In this appendix, we prove that an ω-complete and complete theory Γ derives a contradiction.

Lemma 1.(Gödel 1931) *Let* Γ *be any given theory of arithmetic that can represent all primitive recursive functions. For each formula* $\varphi(x)$ *with one variable x, there exists a sentence* ψ *such that*

$$\Gamma \vdash \varphi(\overline{\ulcorner\psi\urcorner}) \equiv \psi.$$

Proof. The proof is in Gödel(1931:173-177).[13)]

We consider a provability predicate $Prf_{\Gamma}(x,y)$ which expresses that x is a derivation of y from the axioms of Γ.[14)] $Prf_{\Gamma}(x,y)$ satisfies the following relations:

$$\Gamma \vdash \varphi \text{ iff for some } n,\ Prf_{\Gamma}(\ulcorner\varphi\urcorner, n) \text{ is provable in } \Gamma \quad (1)$$

$$\Gamma \nvdash \varphi \text{ iff for all } n,\ {\sim}Prf_{\Gamma}(\ulcorner\varphi\urcorner, n) \text{ is provable in } \Gamma \quad (2)$$

Applying Lemma 1 to ${\sim}\exists y Prf_{\Gamma}(x,y)$, *we have*

---

13) Lemma 1 is due to Gödel(1931) and often called 'fixpoint lemma' or 'diagonal lemma.' He does not state the lemma explicitly, but his proof of Theorem VI in Gödel(1931) includes it.

14) It is a well-known fact that a provability predicate $Prf_{\Gamma}(x,y)$ can be recursively defined. A precise formulation of $Prf_{\Gamma}(x,y)$ appears in Gödel (1931, p. 171). Gödel's '*xBy*' has the same role of $Prf_{\Gamma}(x,y)$.

$$\Gamma \vdash \varphi \equiv \sim\exists y \overline{Prf_{\Gamma}}(\overline{\ulcorner\varphi\urcorner}, y). \tag{3}$$

Having (1), (2) and (3), we have a contradiction true in an ω-complete and complete theory Γ.

Theorem 1. *Let Γ be any given ω-complete theory of arithmetic that can represent all primitive recursive functions. If Γ is complete, there exists φ such that*

$$\Gamma \vdash \varphi \wedge \sim\varphi.$$

Proof. To prove $\Gamma \vdash \varphi \wedge \sim\varphi$, we have three claims.

Claim 1. $\Gamma \nvdash \varphi$.

Suppose $\Gamma \vdash \varphi$. By (1), there is an n such that $Prf_{\Gamma}(\ulcorner\varphi\urcorner, n)$, so $\Gamma \vdash \exists y \overline{Prf_{\Gamma}}\left(\overline{\ulcorner\varphi\urcorner}, y\right)$. Since we have (3), $\Gamma \vdash \sim\exists y \overline{Prv_{T}}\left(\overline{\ulcorner\varphi\urcorner}, y\right)$ which contradicts $\Gamma \vdash \exists y \overline{Prf_{\Gamma}}\left(\overline{\ulcorner\varphi\urcorner}, y\right)$. It is not the case that $\Gamma \vdash \varphi$. Therefore, $\Gamma \nvdash \varphi$.

Claim 2. $\Gamma \nvdash \sim\varphi$.

Suppose $\Gamma \vdash \sim\varphi$. As we have $\Gamma \vdash \sim\varphi$ and (3), classically $\Gamma \vdash \exists y \overline{Prf_{\Gamma}}\left(\overline{\ulcorner\varphi\urcorner}, y\right)$. Since Γ is ω-complete, there exists n such that $Prf_{\Gamma}(\ulcorner\varphi\urcorner, n)$. Hence, by (1), $\Gamma \vdash \varphi$ which contradicts the supposition. Therefore, $\Gamma \nvdash \sim\varphi$.

Claim 3. If Γ is complete, $\Gamma \vdash \varphi \wedge \sim\varphi$.

Suppose Γ is complete. $\Gamma \nvdash \varphi$ implies $\Gamma \vdash \sim\varphi$ and $\Gamma \nvdash \sim\varphi$ classically implies $\Gamma \vdash \varphi$. From the claim 1 and 2, we have $\Gamma \vdash \varphi \wedge \sim\varphi$. Q.E.D.

Appendix B. Deriving a Contradiction from Rosser's Proof.

Barkley Rosser (1936) showed that consistency could replace ω-consistency in Gödel's incompleteness theorem. For the purpose of deriving a contradiction, we introduce additional terminologies and arithmetical facts.

Let define $x < y$ as $\exists z(x + S(z) = y)$ where $S(x)$ is a successor function and $x \leq y$ as $(x < y) \vee (x = y)$. We have the following fact from the theory Γ of *PA*.

$$\Gamma \vdash \forall y(y < \overline{m} \vee y = \overline{m} \vee \overline{m} < y) \qquad (4)$$

We define a Rosser's predicate $R_\Gamma(\overline{\lceil x \rceil})$ as $\forall y(Prf_\Gamma(x,y) \supset \exists_{z<y}(Prf_\Gamma(neg(x),z)))$ with $neg(\lceil \psi \rceil) = \lceil \sim \psi \rceil$. As $Prf_\Gamma(x,y)$ is, $\exists_{z<y}(Prf_\Gamma(neg(x),z))$ is recursively defined. An application of Lemma 1 of Appendix A yields a ψ such that

$$\Gamma \vdash R_\Gamma(\overline{\lceil \psi \rceil}) \equiv \psi \qquad (5)$$

Theorem A shows that $\Gamma \vdash \psi \equiv \sim \psi$ if Γ is consistent.

Theorem A. *Let Γ be a theory of arithmetic that can represent all primitive recursive functions. If Γ is consistent, there exists ψ such that*

$$\Gamma \vdash \psi \equiv \sim \psi.$$

Proof. Let Γ be consistent and ψ be a sentence satisfying (5).

Claim 1. $\Gamma \vdash \psi \supset \sim \psi$.

Suppose $\Gamma \vdash \psi$. By (1), there exists an n such that $Prf_\Gamma(\lceil \psi \rceil, n)$. Since Γ is consistent, no m satisfies $Prf_\Gamma(\lceil \sim\psi \rceil, y)$. We have, for some $\overline{n}$,

$$\Gamma \vdash \overline{Prf_\Gamma}\left(\overline{\lceil \psi \rceil}, \overline{n}\right) \wedge \sim\exists_{z<\overline{n}}\left(\overline{Prf_\Gamma}\left(\overline{\lceil \sim\psi \rceil}, z\right)\right) \qquad (6)$$

From (6), we use classical logic and derive the following:

$$\Gamma \vdash \sim\forall y(\overline{Prf_\Gamma}\left(\overline{\lceil \psi \rceil}, y\right) \supset \exists_{z<y}\left(\overline{Prf_\Gamma}\left(\overline{\lceil \sim\psi \rceil}, z\right)\right)) \qquad (7)$$

(7) means $\Gamma \vdash \sim\overline{R_\Gamma}(\overline{\lceil \psi \rceil})$. Applying $\sim R_\Gamma(\lceil \psi \rceil)$ to (2), we have

$$\Gamma \vdash {\sim}\psi.$$

Therefore, it follows that

$$\Gamma \vdash \psi \supset {\sim}\, \psi.$$

Claim 2. $\Gamma \vdash {\sim}\psi \supset \psi$.

Suppose $\Gamma \vdash {\sim}\psi$. *For some m,* $Prf_{\Gamma}(\ulcorner {\sim}\psi \urcorner, m)$. *Since* Γ *is consistent, for all n,* ${\sim}Prf_{\Gamma}(\ulcorner \psi \urcorner, n)$ *and in particular, for all* $n \leq m$ *does so. We have*

$$\Gamma \vdash \forall y((y < \overline{m} \vee y = \overline{m}) \supset {\sim}\overline{Prf_{\Gamma}}\left(\overline{\ulcorner \psi \urcorner}, y\right)) \qquad (8)$$

From (4) and (8), it follows that

$$\Gamma \vdash \forall y(\overline{Prf_{\Gamma}}\left(\overline{\ulcorner \psi \urcorner}, y\right) \supset (\overline{m} < y) \qquad (9)$$

In addition, we have

$$\Gamma \vdash \overline{Prf_{\Gamma}}\left(\overline{\ulcorner {\sim}\psi \urcorner}, \overline{m}\right) \qquad (10)$$

It follows from (10) that

$$\Gamma \vdash \forall y(\overline{m} < y \supset \exists_{z<y}(\overline{Prf_{\Gamma}}\left(\overline{\ulcorner {\sim}\psi \urcorner}, \overline{z}\right))) \qquad (11)$$

From (9) and (11), by enthymeme, we have

$$\Gamma \vdash \forall y(\overline{Prf_{\Gamma}}\left(\overline{\ulcorner \psi \urcorner}, y\right) \supset \exists_{z<y}\left(\overline{Prf_{\Gamma}}\left(\overline{\ulcorner {\sim}\psi \urcorner}, z\right)\right)) \qquad (12)$$

Applying (12) to (5), we have

$$\Gamma \vdash \psi.$$

It follows that

$$\Gamma \vdash {\sim}\psi \supset \psi.$$

Therefore, by the claim 1 and 2, $\Gamma \vdash \psi \equiv {\sim}\, \psi$. *Q.E.D.*

It is readily proved that Γ cannot be both consistent and complete, on par with the proofs of Theorem 1 and 2.

Corollary A. (Rosser, 1936). *If Γ is consistent, then Γ is not complete.*

Proof. Let ψ be a sentence satisfying (5).

Claim 1. $\Gamma \nvdash \psi$.

Suppose $\Gamma \vdash \psi$. By Theorem A, we have $\Gamma \vdash \sim\psi$. Contradiction. Hence, $\Gamma \nvdash \psi$.

Claim 2. $\Gamma \nvdash \sim\psi$

Suppose $\Gamma \vdash \sim\psi$. By Theorem A, $\Gamma \vdash \psi$. Hence $\Gamma \nvdash \sim\psi$.

Therefore, Γ is not complete. Q.E.D.

$\Gamma \nvdash \psi$ does not imply $\Gamma \vdash \sim\psi$ if there is no presupposition that Γ is complete. With the completeness presupposition of Γ, we have $\Gamma \vdash \psi \wedge \sim\psi$.

Corollary B. *If Γ is complete, then Γ is not consistent.*

Proof. Let Γ be complete. Then $\Gamma \nvdash \psi$ implies $\Gamma \vdash \sim\psi$ and $\Gamma \nvdash \sim\psi$ classically means $\Gamma \vdash \psi$. From the claim 1 of Corollary A, we have $\Gamma \vdash \sim\psi$ and by Theorem A, $\Gamma \vdash \psi$. Hence, $\Gamma \vdash \psi \wedge \sim\psi$. Q.E.D.

Corollary A and B shows the incompatibility of the consistency and completeness of Γ. In other words, even though we accept Rosser's proof of the incompleteness, Γ cannot have a contradiction without the assumption of completeness. The next thing that Priest has to argue is that any correct theories for arithmetic are complete.

Reference

Anderson, A. R. and Belnap, N. D. (1975), *Entailment*, Princeton New Jersey: Princeton University Press.

Beall, J., Foster, T., and Seligman, J. (2012), "A Note on Freedom from Detachment in the Logic of Paradox", *Notre Dame Journal of Formal Logic*, 54 (1), pp. 15-20.

Chihara, C. S. (1984), "Priest, The Liar, and Gödel", *Journal of Philosophical Logic*, 13, pp. 117-124.

Dummett, M. (1963), "The Philosophical Significance of Gödel's Theorem", in M. Dummett, (ed.), *Truth and Other Enigmas*, Cambridge: Harvard University Press, pp. 186-201.

Gödel, K. (1931), "On formally undecidable propositions of *Principia Mathematica* and related system I", in S. Feferman, J.W. Dawson, S.C. Kleene, G. H. Moore, R.M. Solovay, J. Heijenoort, (eds.), *Kurt Gödel: Collected Works, Volume I*, Oxford: Oxford University Press, pp. 144-195.

Hellman, G. and Bell, J. (2006), "Pluralism and the Foundations of Mathematics", in C.K. Waters et all. (eds.), *Scientific Pluralism*, Minnesota Studies in the Philosophy of Science, Vol. XIX, Minneapolis: University of Minnesota Press, pp. 64-79.

Meyer, R. K. (1976), "Relevant Arithmetic", *Bulletin of the Section of Logic of the Polish Academy of Sciences*, 5, pp. 133-137.

Meyer, R. K. (1996), "Kurt Gödel and the Consistency of $R^{\#}$", in P. Hajek, ed., *Logical Foundations of Mathematics, Computer Science and Physics: Kurt Gödel*'s *Legacy*, Springer, pp. 247-256.

Meyer, R. K. and Mortensen, C. (1984), "Inconsistent Models for Relevant Arithmetics", *The Journal of Symbolic Logic*, 49, pp. 917-929.

Paris, J. B. and Pathmanathan, N. (2006), "A Note on Priest's Finite Inconsistent Arithmetics", *Journal of Philosophical Logic*, 35, pp. 529-537.

Paris, J.B. and Sirokofskich, A. (2008), "On LP-models of Arithmetic", *The Journal of Symbolic Logic*, 73 (1), pp. 212-226.

Priest, G. (1979), "The Logic of Paradox", *Journal of Philosophical Logic*, 8 (1), pp. 219-241.

Priest, G. (1984), "Logic of Paradox Revisited", *Journal of Philosophical Logic*, 13, pp. 153-179.

Priest, G. (1991), "Minimally inconsistent LP", *Studia Logica*, 50, pp. 321-331.

Priest, G. (1997), "Inconsistent Models for Arithmetic: I, Finite Models", *The Journal of Philosophical Logic*, 26, pp. 223-235.

Priest, G. (2000), "Inconsistent Models for Arithmetic: II, The General Case", *The Journal of Symbolic Logic*, 65, pp. 1519-1529.

Priest, G. (2005), *Towards Non-Being*, Oxford University Press.

Priest, G. (2006a), *In Contradiction: A Study of the Transconsistent* (expanded ed.), Clarendon: Oxford University Press.

Priest, G. (2006b), *Doubt Truth to be a Liar*, Clarendon: Oxford University Press.

Priest, G. (2013), "Mathematical Pluralism", Logic Journal of IGPL, 21 (1), pp. 4-13.

Shapiro, S. (2002), "Incompleteness and Inconsistency", *Mind*, 111, pp. 817-832.

Rosser, R. (1936), "Extensions of Some Theorems of Gödel and Church", *The Journal of Symbolic Logic*, 1, pp. 87-91.

Tarski, A. (1933), "Some Observations on the Concepts of ω-consistency and ω-completeness", in J. Corcoran, *Logic, Semantics, Metamathematics: Papers from 1922 to 1938*, 2nd ed., Hackett Publishing Company, pp. 279-295.

Tennant, N. (2004), "Anti-Realist Critique of Dialetheism", in G. Priest, JC Beall, and B.A. Garb, Eds., *The Law of Non-Contradiction*, Clarendon: Oxford University Press, pp. 355-384.

고려대학교 철학과

The Department of Philosophy, Korea University

choi.seungrak.eddy@gmail.com

괴델의 불완전성 정리가 양진주의의 근거가 될 수 있는가?

최 승 락

양진주의는 참인 모순이 존재한다는 입장이다. 필자는 이 글에서 괴델 정리가 양진주의의 근거라는 프리스트의 논변이 설득력이 없음을 논할 것이다. 이는 괴델 증명이 우리에게 주는 교훈은 임의의 충분히 강한 산수에 관한 이론이 완전하면서 일관적일 수 없다는 것이기 때문이다. 다음으로 필자는 프리스트의 비일관적이고 완전한 산수에서 모순이 도출될 수 있음을 설명할 것이다. 그리고 괴델 문장이 비일관적이고 완전한 산수이론에 적용되어 양진주의에 관한 대안논변을 제시할 수 있음을 소개하고 이 경우에는 순환성의 문제가 있음을 논할 것이다.

요약해서, 필자는 괴델 정리 및 그와 관련된 정리는 완전한 이론들과 일관적인 이론들 간의 관계를 보여줄 뿐임을 주장할 것이다. 괴델 문장의 적용을 통해 도출된 모순이 중간값과 같은 참인 문장의 값을 지닐 수 있는 것 역시 산수에 관한 비일관 모형에서일 뿐이다. 비일관성이나 완전성에 관한 가정을 하지 않는다면, 괴델 문장의 적용이 참인 모순을 이끌어 낼 수 없으며 그렇기에 괴델 정리 및 그와 관련된 정리는 양진주의의 근거가 될 수 없다.

주요어: 괴델의 불완전성 정리, 로서의 불완전성 정리, 양진주의,
비일관산수, 그래햄 프리스트