Logic and Logical Philosophy
Volume 25 (2016), 203–224
DOI: 10.12775/LLP.2016.012

Marcin Tkaczyk

THE CASE OF DIALETHEISM

Abstract. The concept of dialetheia and the claim of dialetheism has been examined and compared to such related concept as contradiction, antinomy, consistency and paraconsistency. Dialetheia is a true contradiction and dialetheism is the claim that there exists at least one dialetheia. It has been observed that dialetheism is equivalent to the negation of the traditional principle of contradiction. Hence, dialetheism itself is no new idea in whatsoever. The novelty of dialetheism consists in the arguments delivered for its case. Key justification the partisans deliver for dialetheism has been examined and evaluated: antinomies, an alleged Gödel's paradox, and existence of limits of thought. The structure of those arguments has been analyzed. It has been claimed that they share one and the same simple structure which may be called reverse paradox. The vital content dialetheists add to the traditional paradoxes is only the thesis of reliability of the vernacular prima facie knowledge. Three objections have been raised against the justification of dialetheism: firstly, it has been claimed that exactly the same argument supports principle of contradiction, secondly, it has been questioned whether the arguments preserve their value when logic is subject to revision, and thirdly, it has been claimed that the underlying logic of dialetheism is classical.

Keywords: antinomy; dialetheia; dialetheism; consistency; contradiction; paraconsistency; paradox

Introduction

There are theses and there are cases (arguments) for or against the theses. Theses may be true or false (some people think they may also be both or neither true nor false), whereas arguments may be good or bad, strong or weak. The focus of this paper is an argument rather than the

Received March 15, 2016. *Published online* June 7, 2016

thesis itself the argument is supposed to support. Hence, I am not answering the question whether or not the thesis I speak of is true. I rather dare to answer the question whether or not the thesis is justified. The answer is negative, to the effect that the thesis is not justified, and not to the effect that the thesis is not true. And the thesis in question is called dialetheism which is simply the denial of the principle of contradiction.

1. The Concept of Dialetheism

A *dialetheia* is a pair of such sentences that (a) one of those sentences is the negation of the other and (b) both those sentences are true. Having adopted the usual symbol "$\neg$" as the one-place connective of negation, one can say that, for an arbitrary formula φ, a dialetheia is a pair $\varphi, \ulcorner\neg\varphi\urcorner$ of true formulas. Now, *dialetheism* is the thesis that there exists at least one dialetheia, and a *dialetheist* is someone who claims that there exists at least one dialetheia. The related adjective is "dialtetheic". The term "dialetheia" and the related terms have been recently invented in Australia by Graham Priest and Richard Routley [1989, p. XX]. The terms "dialethism" and "dialethic" are admissible versions of the terms "dialetheism" and "dialetheic" respectively [Priest, 2007, p. 131].

Other definitions of dialetheia may be encountered. Sometimes it is a true conjunction $\ulcorner\varphi \wedge \neg\varphi\urcorner$ to be called dialetheia rather that the pair of conjuncts [Priest, 2002a, p. 652]. No confusion may be caused by such ambiguity, provided a conjunction is equivalent to the set of all its conjuncts. Dialetheia may also be defined as a single sentence φ such that both itself and its negation are true. Or even such a sentence φ that it is both true and false.

To understand properly the concept of dialetheia it is necessary to consider its relationship to other contradiction-like concepts and to the principle of contradiction.

Negation in the classical sense is a contradiction-forming connective to the effect that a set Φ of formulas is classically inconsistent if and only if there is such a formula φ that

$$\Phi \vdash \varphi \text{ and } \Phi \vdash \neg\varphi.$$

However, the connective of negation has never been understood as the only way to form inconsistency. For example, in Aristotle's syllogistic the schemata: "a applies to every b" and "a does not apply to some b"

form a contradiction. And there are languages with no negation-like concept at all. Hence, other consistency constraints came into being and the most important one comes from Emil Post: a set of formulas is Post inconsistent if and only if its closure under consequence includes every formula. A classically inconsistent set may also be called *contradictory*, *contradictive* or *antinomial*. A Post inconsistent set is also often called *trivial* or *explosive*.

In classical logic the classical concept of consistency equals to Post's one by means of Duns Scotus' principle (*Explosion*):

$$\text{if both } \Phi \vdash \varphi \text{ and } \Phi \vdash \neg\varphi \text{ for some } \varphi, \text{ then } \Phi \vdash \varphi, \text{ for all } \varphi. \tag{1}$$

The converse relationship is obvious. A stupendously simple proof of Duns Scotus' principle has been delivered as early as XIV century by Pseudo-Scotus (Iohannes de Cornubia):

$$\begin{array}{cl} \varphi, \neg\varphi, & \text{by assumption,} \\ \hline \varphi \vee \psi, \neg\varphi, & \text{by } \varphi \vdash \varphi \vee \psi, \\ \hline \psi, & \text{by } \varphi \vee \psi, \neg\varphi \vdash \psi. \end{array} \tag{2}$$

The proof is clearly based on the two following rules of inference only:

$$\varphi \vdash \varphi \vee \psi, \tag{3a}$$

$$\varphi \vee \psi, \neg\varphi \vdash \psi, \tag{3b}$$

i.e., the principles of *addition* and *disjunctive syllogism* in turn. And the rule itself might have been even accepted by William of Soissons in XII century [Kneale and Kneale, 1962, pp. 201, 282]. Logics with the rule (1) are often called *explosive*, whereas logics that are not explosive are called *paraconsistent* or *non-Scotian*.

The concept of contradiction may cooccur with concepts of proof, justification and others. The cooccurrence of contradiction and proof is called *antinomy*, i.e., a proof of both some formula and its negation in one theory. And there is the whole skeptics' idea of a formula and its negation beeing equally justified as the reason to suspend judgements.

An emphasis should be put on that in all the listed examples the contradiction is considered a challenge, something one should deal with in one way or another. In dialetheism things are completely different. A dialetheia is a cooccurrence of contradiction and truth. Due to the very cooccurrence dialetheists hold that the real world is inconsistent, not only accounts of the world. This is why Bryson Brown calls dialetheists the most radical paraconsistentists who are "indeed against consistency, at least as a global constraint on our metaphysics" [Brown, 2002, p. 628]. Actually, Graham Priest follows Ludwig Wittgenstein to call inconsistency a taboo. Priest counts "the final breaking of the taboo against inconsistency" as a very notable feature of philosophy in the twentieth century [Priest, 2002a, p. 651].

The relation of dialetheism to the principle of contradiction seems to be even more instructive. The principle of contradiction turns out a simple negation of dialetheism, as it forbids the joint affirmation of a sentence and its negation, whereas dialetheism simply allows it. According to the principle of contradiction, for any pair of sentences, if one sentence is the negation of the other, then it is not the case that the sentences are both true. This clearly means exactly that there is no such pair of sentences that (a) one of those sentences is the negation of the other and (b) both those sentences are true. And as such a pair of sentences would be a dialetheia, this means simply that there is no dialetheia. Dialetheism, *au contraire*, means that there exists at least one dialetheia. Hence, the principle of contradiction is simply the negation of dialetheism. A generalised version of the principle of contradiction may be useful to the effect that:

Every set of true sentences is consistent, (4)

whereas dialetheism might be generalised to the thesis:

Some sets of true sentences are not consistent. (5)

The theses (4) and (5) are clearly simple denials to each other.

Notice that although according to dialetheists some sentences are dialetheias, it is not supposed to follow that all sentences are dialetheias. Particularly it is nor supposed to follow that every sentence is true. The thesis that all sentences are true is usually called *trivialism* and rejected by prominent dialetheists. Therefore a dialetheist must subscribe to some paraconsistent (non-Scotian) logic which is not explosive. It is obvious that on the basis oc classical logic dialetheism entails trivialism immediately.

It seems clear that — although the term "dialetheism" is relatively young — the thesis of dialetheism itself is no new stand in logic or metaphysics. The principle of contradiction has been purposefully questioned many times and anawares violated exorbitantly many times. A good deal of such historical analyses may be found in Priest's works [Priest, 2002a,b, 2006, 2007]. And it is Jan Łukasiewicz who challenged principle of contradiction on the fully logical ground more than a century ago [Łukasiewicz, 1987].

The true novelty of dialetheism (except the shapely term) seems to be the way dialetheism is being justified. A number of interesting complex arguments has been invented or analyzed mostly by Graham Priest. The shared schema of such arguments might be called *reverse* paradox (reverse antinomy) argument, as it involves traditional paradoxes (and some new as well) the other way round.

Hundreds of antinomies have been uncovered in advanced theories as well as in the vernacular. The classical routine begins with a set $\varPhi$ of *prima facie* true assumptions. Then a proof is being detected that both

$$\varPhi \vdash \varphi \quad \text{and} \quad \varPhi \vdash \neg\varphi,$$

for some formula φ, meaning simply that the set $\varPhi$ is inconsistent. For the sake of classical logic the *prima facie* affirmation of the assumptions $\varPhi$ gets authomatically suspended and search begins for alternative sets of assumptions, usually — but not necessarily — some weakened versions of $\varPhi$. The routine rests on the assumption (4), which is fundamental for both the classical logic and the classical metaphysics. Simplifying slightly the case, one could say that the routine is an exemplification of the *modus tollens* schema: if my *prima facie* assumption is true then some contradiction is true as well, but no contradiction is true (by the principle of contradiction), therefore my *prima facie* assumption is not true.

A dialetheist reasons rather according the *modus ponens* schema: if my *prima facie* assumption is true then some contradiction is true as well, and my assumption is (obviously) true, therefore some contradiction is true, i.e., there exists a dialetheia. According to a dialetheist it is the thesis (4) which gets under review and eventually turns out to be false under pressure from the antinomy, rather than the *prima facie* true assumptions forming the set $\varPhi$. This is what some authors mean when saying that dialetheists take antinomies at face value. Consider to popular examples of the Liar and Russell's antinomies.

The dialetheist argument template will be examined on three significan examples of traditional self-reference antinomies, an alleged Gödel's paradox, and Priest's favourite idea of *conceptual limits*. In this analysis I follow mostly Priest's recent handbook work [Priest, 2007] with some auxiliary reference to his previous monographic works [Priest, 2002b, 2006].

2. Antinomies of Self-Reference

The famous Liar is presumeably the best known paradox. Traditionally an agent is described admitting to lie at the very moment of admission — provided that "to lie" means simply "to deliver an admission which is not true". If the admission is true, the agent is lying, hence, his admission is not true. If the admission is not true, the agent is not lying, hence, his admission is true. It follows that the admission both is and is not true. Priest refers to a more modern version of the antinomy with the one place predicate "T" of truth and the liar sentence "λ":

$$T(\varphi) \equiv \varphi, \text{ for every } \varphi, \tag{6a}$$

$$\lambda \equiv \neg T(\lambda), \tag{6b}$$

The formula (6a), playing the rôle of definition of truth, makes any sentence φ and the sentence that the φ is true be equivalent. The formula (6b), being a definition of the sentence λ, makes the sentence claim not to be true. By substitution of λ for φ in the row (6a) one obtains that

$$T(\lambda) \equiv \lambda. \tag{6c}$$

The rows (6b) and (6c) entail clearly that

$$T(\lambda) \equiv \neg T(\lambda). \tag{6d}$$

Within the confines of classical propositional calculus there are valid rules of inference for every φ:

$$\begin{aligned} &\varphi \equiv \neg\varphi \vdash \varphi, \\ &\varphi \equiv \neg\varphi \vdash \neg\varphi, \end{aligned} \tag{7}$$

So, from the row (6d) it follows logically both that $T(\lambda)$ and that $\neg T(\lambda)$, meaning that the liar sentence λ both is and is not true. There is a number of solutions to the liar paradox and according to Priest they all fail in one way or another [Priest, 2007, pp. 171–172].

In a series of dilemmas Graham Priest carefully shows that no consistent theory of the truth predicate "T" is liable to express the vernacular concept of truth. [Priest, 2002b, part 3; 2007, p. 174]. Priest claims that "all attempts to solve the paradox swing uncomfortably between inconsistency and a self-refuting inexpressibility" [Priest, 2007, p. 174]. He concludes that:

> The semantic rules that govern notions such as truth over-determine the truth values of some sentences, generating contradictions. The only way

> to avoid this is to dock this richness in some way. But doing this just produces incompleteness, making it the case that it is no longer English that we are talking about. [Priest, 2007, p. 174]

The conclusion is that the vernacular is crucially inconsistent, and yet "perfectly serviceable". According to Priest the inconsistent vernacular serves better that artificial inconsistent languages. The inconsistent common sense theory of truth (and other semantical notions) seems to be preferable to artificial ones. It might be a serious argument in favor of dialetheism: an inconsistent common sense theory seems to deliver the best description of the world.

Another domain of famous antinomies are the foundations of mathematics and the most simple one is Russell's paradox. By the end of the nineteenth century an analysis of numbers in purely logical terms was eagerly searched for. The first mature account of such analysis was delivered by Gottlob Frege who attempted to show that arithmetic was identical with logic. The tool for it was a union of logic and set theory. Frege was the first to show how logic could be developed into arithmetic, provided some theory of extensions of concepts may be itself considered a part of logic [Kneale and Kneale, 1962, pp. 435–436].

A theory of sets (classes) was elaborated as a mathematical discipline by Georg Cantor between 1874 and 1897 basing to some extent on George Boole's algebra of classes. It is assumed that there are two ways to indicate a set: (a) to list its elements, (b) to describe a necessary and sufficient condition for any object to be an element of the set [Kneale and Kneale, 1962, pp. 438–439]. Having identified extensions of concepts with sets Gottlob Frege accepted the idea that a necessary and sufficient condition of membership indicates the set: "I say that something belongs to a class when it falls under the concept whose extension the class is" [quoted in Kneale and Kneale, 1962, p. 653]. As Priest summarizes: "According to Frege's theory of extensions, the simplest and most obvious, every property has an extension" [Priest, 2007, pp. 178–179].

Frege's basic assumption got known as the unrestricted principle of comprehension or of (set) abstraction. Let $\varphi(\alpha)$ be any formula with at most one free variable α. According to the principle of abstraction, the conditions:

x is in the set of such objects x that $\varphi(x)$, (8)
$\varphi(x)$,

are equivalent. In particular x is in the set of natural numbers if and only if x is a natural number, x is in the set of European cities if and only if x is an European city, x is in the set of chimpanzees if and only if x is a chimpanzee, x is in the set of archangels if and only if x is an archangel, x is in the set of dwarfs if and only if x is a dwarf, x is in the set of square circles if and only if x is a square circle, and so forth. It truly may seem obvious that such assumptions are to be legitimately extrapolated for absolutely any meaningful condition. It immediately follows that for any arbitrary condition $\varphi(\alpha)$ with at most one free variable α it is the case that:

$$\exists y \forall x \big(x \in y \equiv \varphi(x)\big), \tag{9a}$$

where "$\in$" is a symbol for the expression "is in" or "is a member of" appearing in the phrase (9a). The principle of abstraction seemed to Frege as logical or obvious, hence, the symbol "$\in$" got considered logical, underlying Frege's analysis of arithmetic as a part of logic.

Unfortunately, the principle of abstraction (9a) turned out to be inconsistent by Bertrand Russel's famous condition "$x \notin x$". In particular, the exemplification:

$$\exists y \forall x (x \in y \equiv x \notin x) \tag{9b}$$

of the principle (9a) is inconsistent. For notice that the formula:

$$\forall x (x \in y \equiv x \notin x) \to (y \in y \equiv y \notin y) \tag{9c}$$

is a theorem of first order logic (and typically even its axiom *dictum de omni*). Therefore, by means of the rules (7) so is the formula:

$$\forall x (x \in y \equiv x \notin x) \to (y \in y \wedge y \notin y).$$

So also:

$$\exists y \forall x (x \in y \equiv x \notin x) \to \exists y (y \in y \wedge y \notin y).$$

Therefore, by *modus tollens*, since "$\neg \exists y (y \in y \wedge y \notin y)$" is a tautology, we obtain the negation of (9b):

$$\neg \exists y \forall x (x \in y \equiv x \notin x),$$

In other words, the antinomy arises in the following way:

$$\exists y \forall x (x \in y \equiv x \notin x) \vdash \exists y (y \in y \equiv y \notin y) \vdash \exists y (y \in y \wedge y \notin y). \tag{9d}$$

It should be emphasised that Frege immedialtely identified the unrestricted principle of abstraction as the source of the antinomy. And yet, it has always been difficult to clearly understand why such a common sense assumption might not be correct [Kneale and Kneale, 1962, p. 653].

A high number of antinomies have been discovered within set theory and enormous effort has been put to cure those antinomies. However, it is to be admitted that no alternative account of the foundations of mathematics invented thus far is based on "the simplest and most obvious" analytic assumptions or could be considered a part of logic itself.

Analogically to the Liar paradox the usual procedure exemplifies to a degree the *modus tollens* schema: As the Principle of Abstraction (9a) entails a contradiction, the principle must be rejected due to the principle of contradiction. Graham Priest exercises the *modus ponens* schema instead: As the Principle of Abstraction (9a) entails a contradiction, and the principle is "the simplest and most obvious" account of extensions, the principle of contradiction must be rejected (i.e., dialetheism must be accepted).

Again, according to Priest, as the common sense set theory is serviceable, one should simply take note of it. There is nothing unexpected in the claim that the vernacular language and the common sense knowledge are replete with contradictions. Priest's idea is that they should be accepted rather than regimented. Among other implications of common sense one should simply take note of that there are dialetheias.

3. Gödel's Paradox

Perhaps the most sophisticated case for dialetheism is the alleged Gödel's paradox. In Graham Priest's opinion arithmetic is provable to be inconsistent by means of Gödel's incompleteness theorem [cf. Priest, 2006, 2007]. Due to a high number of abuses, as well as simple misunderstandings, any philosophical consideration based on Gödel's incompleteness theorem requires an explicit paraphrase of the theorem. Priest's paraphrases are entirely correct and acceptable:

> Gödel showed that in any consistent theory of arithmetic there are sentences such that neither they nor their negations could be proved. [...] To confound matters further, Gödel demonstrated that, given a theory that was intuitively sound, some of the sentences that could not be proved in it could, none the less, be *shown* to be true.
>
> [Priest, 2007, pp. 179–180]

> [...] given any axiomatic and intuitively correct theory of arithmetic, there is a sentence that is not provable in the theory, but which we can yet establish as true by intuitively correct reasoning. The sentence is the famous undecidable sentence that "says of itself that it is not provable"; that is, a sentence γ, of the form $\neg\pi(\langle\gamma\rangle)$.
> [Priest, 2007, p. 186]

Of course, formulas are supposed to be given unique code numbers and $\ulcorner\langle\gamma\rangle\urcorner$ is by definition the numeral of the code number of γ. The number x is the code number of a provable formula if and only if $\pi(x)$. The number $\langle\gamma\rangle$ happens to be the code number of the formula $\neg\pi(\langle\gamma\rangle)$, hence, the formula $\neg\pi(\langle\gamma\rangle)$ equals to γ. So γ is the Gödel's undecidable sentence.

Gödel's theorem is considered by Priest Gödel's paradox. By existence of the true unprovable sentence γ — Priest claims — arithmetic is doomed to inconsistency:

> At any rate, arithmetic is inconsistent, since we can prove certain contradictions to be true; and γ is one of them. [Priest, 2007, p. 186]

According to Priest the antinomy emerges by the fact that it is simultaneously being proved that γ is not provable but true, as proving γ to be true is actually proving γ. By proving γ not to be provable one proves γ. Hence, it is simultaneously being proved that γ is not provable and that γ is provable. The pair γ, $\neg\gamma$ of sentences is to be considered a dialetheia. This is how arithmetic is to turn out to be inconsistent.

Priest offers also a semiformal account of what he calls Gödel's paradox. Let "$\vdash$" be a sign for "our intuitive notion of provability". The argument is based on the following premises:

$$\vdash \pi(\langle\varphi\rangle) \to \varphi, \text{ for every } \varphi, \tag{10a}$$

$$\vdash \neg\pi(\langle\gamma\rangle) \equiv \gamma \tag{10b}$$

The sentence (10a) means that every provable formula is true. According to Priest it is "certainly intuitively correct" and even "analytic". The sentence (10b) defines Gödel's undecidable sentence γ as a sentence saying "of itself that it is not provable". Substituting γ for φ in the sentence (10a) one obtains the sentence:

$$\vdash \pi(\langle\gamma\rangle) \to \gamma,$$

which by the equivalence (10b) entails the sentence:

$$\vdash \pi(\langle\gamma\rangle) \to \neg\pi(\langle\gamma\rangle),$$

which on the base of classical propositional calculus entails obviously that

$$\vdash \neg\pi(\langle\gamma\rangle),$$

and this, again by the equivalence (10b), entails the sentence:

$$\vdash \gamma. \tag{10c}$$

Now the crucial step in Priest's reasoning comes:

> Of course, since we have a proof of γ, we have also demonstrated that $\pi(\langle\gamma\rangle)$. [Priest, 2007, p. 186]

Priest claims clearly that from the sentence (10c) it logically follows that

$$\vdash \pi(\langle\gamma\rangle), \tag{10d}$$

meaning that the code number of the formula γ is a code number of a provable formula. This, by the equivalence (10b), entails the sentence:

$$\vdash \neg\gamma. \tag{10e}$$

The simultaneous proof of the sentence γ and its negation (10c, 10e) should certainly be regarded as an antinomy. And as truth is considered to follow from provability it seems to constitute a dialetheia [Priest, 2007, p. 186].

It seems that the step from (10c) to (10d) is the point of support to Prior's argument. And also Achilles' heel to it. In the recently quoted passage Priest says that proving γ one "of course" authomatically proves also that $\pi(\langle\gamma\rangle)$. But, of course, nothing in this step is a matter of course.

Provability of the formula $\pi(\langle\gamma\rangle)$ in arithmetic follows from the provability of γ itself by the theorem of representation, because the provability relation that $\pi(x)$ within formal theory of arithmetic is established as recursive by means of arithmetization. No proof has been ever delivered to the effect that "our intuitive notion of provability" has a representation in a formal theory of arithmetic. Furthermore, by deriving (10d) form (10c) Priest assumes that a formula φ is "intuitively provable" if and only if it is "intuitively provable" that its code number belongs to the set of formulas which are provable in formal theory of arithmetic, and a formula φ is not "intuitively provable" if and only if it is "intuitively provable" that its code number does not belong to the set of formulas which are provable in some formal theory of arithmetic. Such a risky assumption calls itself for the assumption that "our intuitive notion of provability" is equivalent to the concept of provability within some formal theory. And this is what Priest clearly claims:

> Now consider the canons of mathematical proof, those procedures whereby we establish mathematical claims as true. [...] They are not normally presented axiomatically; they are learned by mathematics students by osmosis. Yet it is reasonable to suppose that they are axiomatic. We are finite creatures; yet we can recognise, in principle, an infinite number of mathematical proofs when we see them. Hence, they must be generated by some finite set of resources. That is, they are axiomatic. [Priest, 2007, p. 186]

This seems to be Leibniz's dream of "mathesis universalis" or "calculemus" revisited. There exists — Prior claims — a formal axiomatic theory exactly equivalent to "our intuitive notion of provability".

Assuming that intuitive arithmetic is axiomatic Priest thinks that $\pi(x)$ if and only if the number x is the code number of a theorem of intuitive arithmetic. Under such condition Priest is legitimate to claim inconsistency of the postulated theory of intuitive arithmetic. On the other hand, according to Priest intuitive arithmetic consists of true sentences only. Therefore, the principle of contradiction turns out to be disproved:

> Now consider the undecidable sentence, γ, for this system of proof. By the theorem, if the system is consistent, we cannot prove γ in it. But — again by the theorem — we can prove γ in an intuitively correct way. Hence, it must be provable in the system, since this encodes our intuitively correct reasoning. By *modus tollens* it follows that the system is inconsistent. Since this system encoded precisely our means of establishing mathematical claims as true, we have a new argument for dialetheism. [Priest, 2007, p. 186]

Priest's reasoning is sophisticated, but its structure turns out to be quite familiar.

It seems arguable that the alleged Gödel's paradox is a stand in a very traditional and well elaborated reasoning. The key observation is that the following four sentences form an inconsistent collection:

every set of true sentences is consistent, (11a)

Φ is a set of true sentences, (11b)

Φ is an axiomatic theory, (11c)

γ is derivable from Φ. (11d)

Of course, $\varPhi$ is an arbitrary set of sentences. The sentence (11a) is simply the version (4) of the principle of contradiction. The sentence (11b) is a general template of Priest's assumption (10a). The sentences (11c) and (11d) create a base for Gödel's theorem. For the set containing all the sentences (11) is not consistent, having accepted any three of the sentences (11), one should authomatically reject the remaining one. It is simply an elementary idea of indirect proof.

Let the set $\varPhi$ contain theorems of arithmetic. Nothing unexpected is there in the set $\varPhi$ being inconsistent. For the sentences (11a) and (11b) give a most obious syllogistic conclusion that $\varPhi$ is consistent, whereas by Gödel's incompleteness theorem the sentences (11c) and (11d) entail inevitably that $\varPhi$ is not consistent. In those circumstances there are four possible outcome conclusions and Priest's reasoning amounts to choosing one of them.

Firstly, if you subscribed to classical logic and arithmetic, you would accept the sentences (11a) and (11b). If you knew also that $\varPhi$ was an axiomatic theory, for example, if $\varPhi$ simply was Peano's theory PA or Robinson's theory Q, you would accept the sentence (11c). From the sentences (11a), (11b) and (11c) you should conclude that γ is not derivable from $\varPhi$, denying the sentence (11d). That would be clearly a version of Godel's first theorem.

Suppose, secondly, you accepted the sentences (11a) and (11b). However, you were not certain of $\varPhi$ being an axiomatic theory, but you derived the sentence γ from $\varPhi$ — for example, if $\varGamma$ happened to be Priest's intuitive arithmetic — you would accept the sentence (11d). If it was the case, you should infer that $\varPhi$ is not an axiomatic theory, denying the sentence (11c). This is also what most logicians do. They do not suppose intuitive arithmetic be axiomatic (actually some of them doubt it be at all a theory). This is why they often think of Gödel's theorem and of similar theorems as describing limitations on deductive methods.

Thirdly, Graham Priest does not subscribe to the principle of contradiction, hence, the sentence (11a) seems dubious to him. But he affirms the remaining sentences (11a), (11b), and (11c). The only solution is to conclude from the three sentences (11a), (11b), and (11c), that some sets of true sentences are not consistent, i.e., to deny the principle of contradiction and to affirm dialetheism. Of course a necessary assumption of Priest's argument is that the intuitive arithmetic is an axiomatic theory. Priest's claim it to be a "reasonable" belief. Well, let us agree that it is a matter of faith.

The true fulcrum of Priest's argument is yet the sentence (11b). Without this assumption the argument would be simply to the effect that the intuitive arithmetic is not consistent. Such a conclusion would be metaphysically harmless, and even likely. The last possible solution of the problem of the inconsistency of the sentences (11) would be to question the sentence (11b). The assumption (11b) — or particularly (10a) — plays a significant rôle in some versions of Gödelian proofs (although it is present neither in Gödel's original paper nor in the most sophisticated contemporary accounts of it). It is not questionable as long as Φ contains theorems of a well tested axiomatic theory, like Robinson's or even Peano's arithmetic. Notwithstanding Graham Priest's claim that every intuitive theorem is true seems to be hazardous, and is definitely an example of Priest's general assumption of final reliability of our common sense knowledge.

Hence, there seems to be nothing mysterious in the alleged Gödel's paradox, if anything like this actually exists. According to the Gödel's theorem:

no consistent system of arithmetic is complete.

This is — by Aristotle's syllogistic — equivalent to the claim that

no complete system of arithmetic is consistent.

If Priest's intuitive arithmetic is a complete axiomatic system, it must not be consistent. This is simply Gödel's result. Priest's contribution is the daring claim that every intuitive theorem is true. It immediately follows that there is an inconsistent set of true sentences. All in all we are provided by Priest again with the claim that the vernacular account of the world (in this case the vernacular account of arithmetic) has the final say. And as the vernacular accounts are not consistent, neither is the world. But this is a new paradox of the vernacular and of our common sense — Priest's paradox rather than Gödel's.

A question might be certainly posed, why a dialetheist should be bothered by the fact that (11) is not consistent. A dialetheist could clearly simply take it at face value that all the sentences (11) are intuitively true, and therefore simply true. The indirect proof like arguments I described are reasonable for those subscribing to the classical logic, and especially the principle of contradiction. But why would a dialetheist continue arguing instead of simply observing the inconsistency of the collection (11) of four true statements? I will shortly come back to this question.

4. Limits of Thought

A very characteristic, interesting and (for me) even intriguing argument of Priest's for dialetheism is based on the idea of transcendency or — to use Priest's favourite expression — *limits of thought* (*limits of the mind* and *conceptual limits* are popular alternatives). In such expressions thoughts or concepts are to be understood in an objective sense as contents of human intensional states rather than subjective consciousness [Priest, 2002b, p. 3]. Within the tradition of Polish Logic one could say that it is thoughts or concepts in the logical sense rather than in psychological sense that are involved.

The limits of thought argument is a thoroughbred metaphysics and although it is exemplified by advanced philosophical theories in all periods, the argument itself is amazingly simple. According to Priest, once it is accepted that there are conceptual limits of any sort, dialetheias spring inevitably into existence:

> Such inconsistencies seem to occur, in particular, in the works of those philosophers who argue that there are limits to what can be thought, conceived, decribed. In the very act of theorising, they think, conceive, or describe things that lie beyond the limit. [Priest, 2007, p. 176]

It should be emphasised that it is the very act of imposing some limits of thought that effects in dialetheias and no further deliberate act of violating the limits is required. It does not matter as well what particular kind of conceptual limits is to be involved. It is the idea of the limits not the idea of what the limits are that leads to dialetheias. Priest insists that to speak of some conceptual limits the very limits one speaks of must be crossed:

> [...] such limits are dialetheic; that is, that they are the subject, or locus, of true contradictions. The contradiction, in each case, is simply to the effect that the conceptual processes in question do cross these boundaries. Thus, the limits of thought are boundaries which cannot be crossed, but yet which are crossed. [Priest, 2002b, pp. 3–4]

The automatism of the limits of thought argument is based on the facts that to accept limits of thought involves both (a) to attribute some features to the world that remain beyond the ability of human language or cognition to express or grasp them, and simultaneously (b) to express or grasp those features. Hence, to accept limits of thought means "to say the unsayable". Graham Priest calls the act (a) *clausure* and the act (b) *transcendence* [Priest, 2002b, pp. 3–4, 11]. And any case of the limits of thought argument is thought of by him as an example of the clausure-transcendence pattern.

Priest's account of the limits of thought argument consists of a survey of many historical examples of theories of limits of one kind or another as well as the claim there actually are such limits. Priest distinguishes four kinds of conceptual limits: limits of the expressible, limits of the iterable, limits of cognition, and limits of conception [Priest, 2002b, p. 11]. The limits of the expressible are directly connected to abilities of human language, the limits of the iterable are connected to the idea of mathematical infinity.

Among immense number of examples of conceptual limits Priest lists and analyzes there are three os special interest (I follow the account of Priest's recent handbook work Priest, 2007). Firstly, God appears as the most typical example of transcendence:

> [...] many philosophers have argued that God is literally beyond conception or description. This has not prevented them from saying things about God, though; for example, in explaining why God is beyond conception. [Priest, 2007, p. 176]

If nothing you might know of is God (nothing you might know of is a necessary being), it follows that you cannot know anything of God. However, by claiming that you won clearly some knowledge of God. Secondly, things as they really are according to Immanuel Kant seem to rest beyond the limits of thought:

> Kant espoused the distinction between phenomena (things that can be experienced) and noumena (things that cannot). Our categories of thought apply to the former [...]; but they cannot be applied to the latter [...] In particular, then, one can say nothing about noumena, for to do so would be to apply categories to them. Yet Kant says much about noumena in the *Critique*; he explains, for example, why our categories cannot be applied to them. [Priest, 2007, p. 176]

Since every object you know of is by definition a phenomenon, a question may be posed, how can they know that there are noumena. Thirdly, the logical structure of the world is — according to Wittgenstein — transcendent:

> Propositions express the facts that constitute the world. They can do so because of a commonality of structure. But such structure is not the kind of thing that propositions can be about (for propositions are about objects, and structure is not an object). One can say nothing, therefore, about this structure. Yet the *Tractatus* is largely composed of propositions that describe this structure, and ground the conclusion that it cannot be described. [Priest, 2007, p. 176]

So, everything you might speak of is a matter of fact. But you may speak of the facts, provided they share the logical structure of your language which is no fact. Therefore you can speak of something which is no matter of fact at all.

Graham Priest does not content himself with the historical survey of philosophical theories involving the idea of conceptual limits. He also subscribes to the belief there actually are some objective limits of thought:

> [...] it is without doubt that there are limits to whatever people want to do [...]. [Priest, 2002b, p. 3]
> There certainly are general philosophical reasons for supposing there to be things beyond the limits of thought. [Priest, 2007, p. 177]

The limits in question are based on that: "Finitude is a basic fact of human existence" [Priest, 2002b, p. 3]. For the thesis of conceptual limits is both true and dialetheic, it follows that there are dialetheias.

Although there are many different ontological theories Priest regards as examples of conceptual limits thesis it is actually the very idea of conceptual limits which is the source of dialetheism. For the thesis:

there are conceptual limits (12a)

is both true and inconsistent. The thesis (12a) is true, as human is a finite nature and, by intuition, it would be obviously false to claim that human knowledge or language has no limits. However, on the other hand the thesis (12a) is inconsistent. Inconsistency of the thesis (12a) consists in that the thesis involves both the facet of clausure and the facet of transcendence. There is a facet of clausure in it, as by definition of the limits it is claimed that

one cannot speak of (know of) anything beyond the limits. (12b)

There is also a facet of transcendence in it, as by establishing the limits it is assumed that

there is something beyond the limits. (12c)

In the thesis (12c) one unevitably claims to speak of (know of) something beyond the limits. Therefore, by the thesis (12b) one clearly claims to speak of (know of) something one cannot speak of (know of).

In Priest's account the idea of the limits concerns primarily human knowledge or language. However, it seems arguable that on some more ontological level the thesis (12a) means something like that:

> there exists something different to everything.

This thesis corresponds to the following formula:

$$\exists x \forall y \; x \neq y, \tag{13}$$

which is obviously contradictory, since its negation is a tautology. In other words, in the facet of clausure a universal set is being described, whereas in the facet of transcendence it is being claimed that some objects are not in the set.

Similarly to other arguments for dialetheism Priest's account is mostly the claim that the thesis (12a) is true. It is a *bona-fide* truth and so belongs to the common sense knowledge. Therefore, there is a true inconsistent sentence. And similarly philosophers have been aware of inconsistency of the unrestricted thesis (12a) for centuries. To avoid the inconsistency philosophers developed a number of theories to have ontology be consistent. Aristotle was the first who claimed that the terms "being" and "existence" are ambiguous. Generally the scope of quantifiers in the formula (13) has been restricted in many ways. Priest, *au contraire*, prefers the common sense "raw", unprocessed thesis (12a) to sophisticated ontological restriction. As in other arguments he simply involves a *modus ponens* inference instead of traditional *modus tollens*.

5. The Essence of Dialetheism

Graham Priest's cases for dialetheism, however sophisticated they are in detail, share one general structure. Consider a set Φ of *prima facie* true sentences and suppose the set turns out to be inconsistent. A partisan of classical logic accepts the fact that some members of Φ must be rejected as false and searches for a weaker but consistent set Φ' to the effect that

$$C(\Phi') \subset C(\Phi),$$

but Φ' is theoretically serviceable and rejection of the respective sentences is justifiable. So to say, if you encounter inconsistency, search for an error in assumptions. Some general idea of *modus tollens* is here clearly involved.

A dialetheist, quite the reverse, would search for such a set Φ' of *bona-fide* truths that

$$C(\Phi) \subseteq C(\Phi'),$$

i.e., a non weaker set Φ' of reliable sentences. And from the existence of a serviceable and reliable, but inconsistent, set Φ' a dialetheist would conclude that the principle of contradiction is false.

Although the a dialetheist's arguments are interesting, I can see at least three serious general objections to be raised against such arguments. Firstly, the key assumption of reliability of the vernacular supports classical logic. Secondly, once logic is subject to revision so are all *prima facie* contradictions, including alleged dialetheias. Thirdly, willy-nilly, when arguing for dialetheism a dialetheist involves and assumes classical logic, including the principle of contradiction.

The first objection concerns the content of the common sense knowledge. The *bona-fide* truths a dialetheist relies on belong mostly to the common sense knowledge. For example, the unrestricted definition of truth (6a), the unrestricted principle of abstraction (9a) are *bona-fide* truths, and so are assumptions that there exist axioms of intuitive arithmetic and limits of thought. As those assumptions are *bona-fide* truths and entail contradictions, it should be accepted that there exist dialetheias.

It is arguable that there is actually one crucial assumption Graham Priest makes, and all the arguments for dialetheism he delivers are actually this very assumption in different disguises. The assumption is that

> the vernacular (common sense, intuitive) knowledge is reliable

(at least until the knowledge is serviceable). This is exactly what Priest wants readers to accept. As the intuitive knowledge is filled with contradiction, Priest's advice is to learn to live with it instead of searching how to cure it.

If it is correct, Priest certainly mislead his case, since the principle of contradiction and the whole classical logic is not less serviceable than any paraconsistent logic (imagine the building of sciences, built on the classical assumptions). And classical logic is definitely miles more vernacular, common sense and intuitive than the jungle of paraconsistent calculi. Return to the principle of explosion (1) which is unacceptable for a dialetheist. The proof (2) of the principle rests on two dazzling vernacular assumptions exclusively, i.e., the rule of addition (3a) and the rule of disjunctive syllogism (3b). No fair argument has been delivered by Priest to the effect that the most simple assumptions of classical logic are less intuitive than his claims of unrestricted definition of truth, unrestricted principle of abstraction, axiomatiziability of the whole intuitive arithmetic etc.

I can agree with Priest willingly that there is no decisive argument for the principle of contradiction (at list I wish I had one). But dialetheism is no more justified, not even by a hairbreadth. Priest's appeal to intuition and the vernacular knowledge form at least equally good basis for classical logic as for dialetheism. Our intuitions and the vernacular are simply inconsistent. This is what we know. Whether to accept it or to cure it remains a matter of choice.

The second objection concerns the proper scope of demanded revision of logic. Once dialetheism has been chosen, it demands a revision of logic. If dialetheism was accepted together with classical logic, every sentence should be considered true. Classical logic is explosive since it contains Duns Scotus Rule (1). To avoid such a conclusion — which is known as trivialism — a dialetheist must subscribe to a paraconsistent (non-Scotian) logic. At least two vital problems arise here.

First, all the antinomies have been constructed within the confines of classical logic. As any paraconsistent calculus is weaker than the classical one it should be answered whether or not *prima facie* dialetheias are to be regarded as dialetheias within the confines of the new logic. Dialetheism may clearly turn out to be self-annihilating.

Second, there is wide range of quite different paraconsistent calculi [see e.g. Brown, 2002]. If some revision of logic is inescapable, it might be an option to search purposefully for a logic pursuing two aims: (a) to be paraconsistent and (b) to avoid antinomies. For example, if Graham Priest is allowed to reject the rules (3a) and (3b), I may claim the right to reject the rules (7) or simply the rule:

$$p \equiv \neg p \;\vdash\; p \wedge \neg p. \tag{14}$$

Those rules are not more established by intuition than the rules (3a) and (3b). And yet, notice that the rules (14) or (7) play a significant rôle in typical antinomies. Formulas like (6d) or the consequent of the formula (9c) create antinomies because of the rules (7) or (14). An example of such inference has been delivered in the row (9d).

Generally, logic is a vital element of any antinomy (or paradox). Once this logic is subject to revision, the antinomy no longer exists and should be created again (or not) within the confines of a new logic. But this logic has not been chosen yet. And if a dialetheia is legitimate to demand it to be non-explosive, I claim my equal right to demand it to be non-antinomial or non-dialetheic. Again, no piece of fair argument for dialetheism has been delivered.

The final objection against dialetheists' arguments is that those arguments themselves assume classical logic, and especially the principle of contradiction.

We have observed this fact already when discussing inconsistency of the collection (11) of sentences. The question is, if a dialetheist was honest, why would he felt commited to reject the sentence (11a) under the pain of inconsistency of the collection (11) of sentences. Applying Priest's argument template to the collection (11) consequently, one should accept the principle of contradiction as a part of intuitive arithmetic and then agree that the arithmetic both is and is not consistent (another beautiful dialetheia).

Similarly for other arguments. Once dialetheism is seriously taken, there is no reason to reject the principle of contradiction, the Duns Scotus Rule (Explosion) or classical logic any more. An honest dialetheist would rather face the truth that all sentences are true and are not, the principle of contradiction holds and does not hold etc. However, Graham Priest supports dialetheism in the object language and fully subscribes to classical logic in the metalanguage. Hence, by a classical rule dialetheists accept:

$$p \to \neg p \vdash \neg p,$$

as the principle of contradiction is being actually assumed by dialetheists' arguments, the thesis of dialetheism should be rejected by self-annihilation.

References

Brown, B., 2002, "On paraconsistency", pages 628–650 in: *A Companion to Philosophical Logic*, D. Jacquette (ed.), Blackwell, Oxford. DOI:10.1002/9780470996751.ch40

Kneale, W., and M. Kneale, 1962, *The Development of Logic*, Clarendon Press, Oxford.

Łukasiewicz, J., 1987, *O zasadzie sprzeczności u Arystotelesa* (On the Principle of Contradiction in Aristotle), PWN, Warszawa. See also: J. Łukasiewicz, "On the principle of contradiction in Aristotle", V. Wedin (trans.), *The Review of Metaphysics*, 24, 3 (1971): 485–509.

Priest, G., and R. Routley, 1989, "The philosophical significance and inevitability of paraconsistency", in *Paraconsistent Logics: Essays on the Inconsistent*, G. Priest, R. Routley, J. Norman (eds.), Philosophia Verlag, Münich.

Priest, G., 2002a, "Logicians setting together contradictories: A perspective of relevance, paraconsistency, and dialetheism", pages 651–664 in *A Companion to Philosophical Logic*, D. Jacquette (ed.), Blackwell, Oxford. DOI:10.1002/9780470996751.ch41

Priest, G., 2002b, *Beyond the Limits of Thought*, Oxford Clarendon Press, New York. DOI:10.1093/acprof:oso/9780199254057.001.0001

Priest, G., 2006, *In Contradiction: A Study of the Transconsistent*, Oxford University Press.

Priest, G., 2007, "Paraconsistency and dialetheism", pages 129–204 in *Handbook of the History of Logic*, vol. 8, D. M. Gabbay, J. Woods (eds.), North Holland, Amsterdam. DOI:10.1016/S1874-5857(07)80006-9

Marcin Tkaczyk
Faculty of Philosophy
Departament of Logic
The John Paul II Catholic University of Lublin
al. Racławickie 14
20-950 Lublin, Poland
tkaczyk@kul.pl