Australasian Journal of Philosophy
Vol. 90, No. 4, pp. 789–795; December 2012

Routledge
Taylor & Francis Group

DISCUSSION NOTE

DEFINITION INCLOSED: A REPLY TO ZHONG

Graham Priest

In 'Definability and the Structure of Logical Paradoxes' (*Australasian Journal of Philosophy*, this issue) Haixia Zhong takes issue with an account of the paradoxes of self-reference to be found in *Beyond the Limits of Thought* [Priest 1995]. The point of this note is to explain why the critique does not succeed. The criterion for distinguishing between the set-theoretic and the semantic paradoxes offered does not get the division right; the semantic paradoxes are not given a uniform solution; no reason is provided as to why the naïve denotation relation is 'indefinite' (other than that its definiteness leads to contradiction); and the account of the denotation relation given clearly misses the mark, even by consistent standards.

Keywords: Inclosure Schema, definability, Berry's Paradox, Graham Priest

1. Introduction: the Issue

In 'Definability and the Structure of Logical Paradoxes'[1] Haixia Zhong takes issue with an account of the paradoxes of self-reference to be found in *Beyond the Limits of Thought*.[2] The argument is executed with an admirable technical deftness; and there is certainly something philosophically attractive about Zhong's main idea. But in the end, I think, it does not work. The point of this note is to explain why.

According to the account given in *BLoT*, all the standard paradoxes of self-reference, set-theoretic and semantic, fit the Inclosure Schema, which is as follows. There are monadic predicates, φ and ψ, and a monadic function, δ, such that it appears to be the case that, where $\Omega = \{y\colon \varphi(y)\}$:

Ω exists and $\psi(\Omega)$
If $x \subseteq \Omega$ and $\psi(x)$:
$\delta(x) \notin x$ [Transcendence]
$\delta(x) \in \Omega$ [Closure]

The contradiction arises when x is Ω. For then $\delta(\Omega) \in \Omega$ and $\delta(\Omega) \notin \Omega$. Since all the paradoxes are generated in the same way, they should all have the

[1]Zhong [2012]. Page and section references are to this.
[2]Priest [1995]. Hereafter, *BLoT*. See esp. Part 3.

Australasian Journal of Philosophy
ISSN 0004-8402 print/ISSN 1471-6828 online © 2012 Australasian Association of Philosophy
http://www.tandf.co.uk/journals http://dx.doi.org/10.1080/00048402.2012.683014

same solution (the Principle of Uniform Solution).[3] *BLoT* advocates a dialetheic solution.

According to Zhong, although all the paradoxes fit the Schema, we may distinguish between the case where the condition $\psi(x)$ is the vacuous condition, $x = x$, and the case in which it is a non-vacuous condition. The first case gives the set-theoretic paradoxes; the second gives the semantic paradoxes.[4] In both, Ω does not exist. In the first, because the set is too large; in the second, because its specification employs an 'indefinite' notion, such as 'is definable'. In both cases, then, the contradiction fails to arise, but for different reasons; hence there are two kinds of paradox, with two different solutions.[5]

2. Vacuity and the Two Kinds of Paradox

A number of things make this position indefensible. First of all, note that there are inclosure paradoxes which use intentional notions, such as 'conceive', 'think'.[6] For example, in Berkeley's paradox, concerning an object which is not being conceived, $\psi(x)$ is 'x is conceivable'.[7] Another example: Kripke has recently published a paradox of self-reference concerning the set of all times when I am thinking about a set of times that does not contain that time.[8] In this, $\psi(x)$ is $\exists t \forall y(\theta_t(y) \leftrightarrow y = x)$, where the particular quantifier quantifies over times, and $\theta_t(y)$ is 'I am thinking of y at time t.[9]

In her paper, Zhong does not mention these paradoxes, so it is not clear what she wants to say about them. The most natural thing, it seems to me, is to assimilate them into the semantic paradoxes. Thus, one may suppose that to conceive/think of x is to bring before the mind a name that refers to x (that is, which defines it in the relevant sense).[10] And given that they are semantic paradoxes, they should have the same sort of solution as the more usual kind, by the Principle of Uniform Solution.

[3]Zhong says (n. 2) that the point of advocating a common structure to the paradoxes is to support the Principle of Uniform Solution. In fact, it is to justify *applying* the Principle to the family of paradoxes of self-reference.

[4]'Why couldn't the difference between semantic and set-theoretic paradoxes exactly consist of this new property $\psi(x)$?' (§3).

[5]The first statement of the Inclosure Schema is by Russell [1907]. His concern there is with the set-theoretic paradoxes, and he states the schema only, in effect, with the vacuous condition, which is therefore omitted. However, by Russell [1908] the view has morphed into the Vicious Circle Principle, which applies explicitly to all the paradoxes. The Vicious Circle Principle is, in effect, a version of the Inclosure Schema. Let $\varphi(y)$ be 'y is a set of any order', and $\psi(x)$ be 'x contains the sets of one particular order' (and so is non-vacuous). If $\psi(x)$, then any specification of a set which quantifies over x is a diagonalizer, $\delta(x)$: the specification defines a set of a higher order, and so one not in x; but it is still a perfectly good set of some order. Of course, the Schema cannot be stated in the official language of the theory of orders. But, as is often observed, though Russell stated and endorsed the Vicious Circle Principle, this cannot be stated in the theory of orders. (See e.g., *BLoT*, 9.6, 9.7.) So much the worse for the theory of orders. At any rate, though Russell may never have stated the Inclosure Schema explicitly in the general form, it certainly informs his mature theory of all the paradoxes of self-reference. It is therefore a little misleading to suggest, as Zhong does, that the formulation of the Inclosure Schema to apply to the semantic paradoxes is not Russell's. ('Priest ... wants to extend Russell's idea to the semantic paradoxes' (§2).)

[6]On intentional paradoxes, see Priest [1991].

[7]See *BLoT*, 9.4.

[8]Kripke [2011].

[9]See Priest [2011].

[10]See *BLoT*, 4.8.

But whatever one says about this point,[11] there are paradoxes in the semantic family in which the condition $\psi(x)$ is vacuous. Thus, in the formulation of the Liar paradox where truth-bearers are taken to be sentences, formulating the appropriate liar sentence requires the set in question to be definable. Thus, given a set, x, we need a sentence which says of itself that it is not in x. This requires x to have a name. However, if truth-bearers are taken to be propositions, this is unnecessary. For *every* set, x, there is a proposition of the form 'This proposition is not in x'.[12] One might suggest that this version of the paradox is really a set-theoretic paradox, and is solved by the fact that there is no set of all propositions (and thus no self-referential proposition when $x = \Omega$). But to force some versions of the liar paradox into one family and some into another would seem a desperate measure.

And even this will not do. There is a paradox which *BLoT* calls the *5th Antinomy*. Starting with any object, there is a thought of it, a thought of the thought of it, a thought of the thought of the thought of it, and so on. We can iterate this procedure into the transfinite, at limit ordinals taking the thought of all the things so far generated. There can be no thought of the whole totality. If there were, it would be the next member of the sequence; by construction, there is no such thing. But there is such a thought; you have just had it. In the 5th Antinomy, $\psi(x)$ is the vacuous condition.[13] Assuming that the appearance of the intentional notion locates this paradox in the semantic family, again we see that in a semantic paradox $\psi(x)$ can be vacuous.[14]

Once one gets away from the myopia induced by focusing on just the usual suspects, the boundary between the set-theoretic and the semantic paradoxes is, in fact, very messy, and there would seem to be no good way of disentangling the two.

3. Definition

Next, according to Zhong, in the case of the semantic paradoxes, Ω does not exist because the notion used to define it is not well-behaved. In some of the semantic paradoxes the condition in question concerns definability; in others it concerns truth, knowability, and (if we are right about the intentional paradoxes) thought, conceivability.[15] If, then, the semantic paradoxes are to have a uniform solution, there must be a uniform story to be told about the misbehaviour of such notions. Zhong seems to agree: 'As Zermelo … has pointed out, the problem with the semantic paradoxes is that the defining criterion which aims to define the totality is not definite, for there are

[11]In correspondence, Zhong says that she is inclined to agree with it.
[12]See *BLoT*, 10.2.
[13]See *BLoT*, 9.2.
[14]Indeed, whether one classifies the paradox as set-theoretic or semantic, Zhong's account is in trouble. For her, in both cases, the solution is that the relevant Ω does not exist. But, notoriously, one can think of non-existent objects—even when their specification has a certain indeterminacy, e.g., 'what Sherlock Holmes had for breakfast on the day he died'.
[15]See the tables in *BLoT*, 9.4 and 10.2.

semantic notions involved and the criterion for whether or not a semantic notion is applicable to a given object is indefinite' (§6). However, she does *not* give a uniform solution to the semantic paradoxes. At the end of §5, Zhong appears to endorse a Tarskian account of truth.[16] This is of a kind different from the account of definability she advances, as we will see in a moment.[17]

Why, though, is the notion of definability not well-defined? Zhong (§4) points out the following. In his axiomatization of set theory, Zermelo required the conditions used to define sets to be 'determinate'. He did not specify what this meant. Later, Skolem characterized it as being a condition expressible in a first-order language with only the predicates '$=$' and '$\in$'. The notion of definability cannot be expressed in this language. So, says Zhong, it is not definite. This reasoning can hardly be correct. Being expressible in the way Skolem suggested may be adequate for pure set-theory. But there are many perfectly good determinate (whatever that means) predicates which cannot be so expressed: 'is an electron', 'weighs exactly *n* grams', 'is midday January 1st, 2012, GMT'. (Note, in particular, that these are not vague predicates.) Why isn't definability like that?

In her next section Zhong contrasts definability as a notion expressed in natural language (ill-defined) with definability as expressed in a formal language (well-defined). When one takes definability to be determined by 'the formal language standard, paradox vanishes' (§5). (Why this is supposed to be so, we will see in a moment.) The contrast is not a good one, though. Maybe natural language notions are less clear than formal notions—though as far as that goes, there seems no particularly good reason to suppose that 'is definable' is in worse shape than 'not' or 'if'. But there are many formal languages with precise proof procedures, etc., that may be taken to formalize the notion of definability. We will come to the one Zhong favours in a moment. This is consistent. But there are inconsistent ones which are no less precise and rigorous, such as the one given in the formalization of Berry's paradox to be found in Priest [1987, 1.8]. The question is only: which is the formalization that is most faithful to the naïve (natural language) notion?

4. Provability and Truth

Which brings us to Zhong's preferred formalization. Zhong defines a notion of definability as follows (§5). Given an axiomatic theory, **T**, which contains standard arithmetic:

(*) A number n is *denominated* in **T** by a formula $\varphi(x)$ iff '$\forall x(\varphi(x) \rightarrow x = \mathbf{n})$' is provable in **T**

[16]'As Tarski ... has pointed out, the effective way to block the Liar paradox is to distinguish between object language and metalanguage' (§5).
[17]In correspondence Zhong informs me that her words should not be taken as an endorsement of a Tarskian solution; she prefers a Kripkean solution. The point remains.

where '**n**' is the numeral for the number n. This is naturally understood as a definition in the metatheory of **T**. However, by standard techniques, the notion of a formula of **T** and provability in **T** can be expressed in **T**. Thus, 'n is denominated by formula $\varphi(x)$' can be coded up as a formula of **T**, $Den(\mathbf{n},\varphi)$.[18] Using the denotation (denomination) predicate, one can express an analogue of the Berry condition 'is the least number not denoted in less than $10 \uparrow 10$ symbols'. This is satisfied by some number, m. But, as Zhong observes, if **T** is consistent:

(#) **m** is the least number not denoted in less than $10 \uparrow 10$ symbols

cannot be proved in **T** (and neither can its negation, since this is false in the standard model). Berry's paradox is thus turned into a formal undecidability result, in the same way that Gödel turned a paradox of self-reference into a formal undecidability result.

Of course, if **T** is inconsistent, the sentence may well be provable, in the same way that an inconsistent theory can prove its own "Gödel undecidable sentence".[19] If this is an argument against dialetheism, then, Zhong's simple *assumption* that **T** is consistent begs the question. However, there are more fundamental things to be said about Zhong's approach.

Return to the definition (*). This does not define what it is for a number to be defined; it defines what it is for a number to be *provably* defined. And provable definability is no more definability than provable truth is truth. (Both (#) and its negation are unprovable, but at least one of them is true.) Now, there is certainly a paradox about the least number not provably defined in less than such and such number of words (symbols),[20] but it is not Berry's paradox, any more than the paradox concerning the sentence 'This sentence is not provable (provably true)' is the liar paradox. As Zhong herself notes, though her formal construction 'is supposed to be a precise formal regimentation of the notion "definable", it cannot fully capture all the intuitions involved in the latter' (§6). Of course not, since it interpolates the notion of provability into the definition. Indeed, as Zhong herself points out (in the text to n. 9), there *is* a condition that defines m in the standard sense. In the footnote, she says that this is only a loose sense. It is not: it is the precise model-theoretic sense.

The correct definition of denomination (definability) is the metatheoretic statement which replaces 'is provable' in (*) with 'is true', thus:

(**) A number n is *denominated* in **T** by a formula $\varphi(x)$ iff '$\forall x(\varphi(x) \rightarrow x = \mathbf{n})$' is true in the intended interpretation of **T**.

[18]More correctly, the second argument here should be $<\varphi>$ where this is the code of φ. However, Zhong appears to identify formulas with their codes. This is harmless if correctly understood.

[19]See the second edition of Priest [1987, ch. 17].

[20]There is a least number not definable in $10 \uparrow 10$ symbols. *A fortiori*, there is a least number not provably definable in $10 \uparrow 10$ symbols. But 'the least number not definable in $10 \uparrow 10$ symbols' (which does have less than $10 \uparrow 10$ symbols) does define it; and provably so, given standard principles concerning descriptions and denotation. We might call this 'Zhong's paradox'.

This statement cannot be coded in arithmetic unless the arithmetic has a truth predicate, which it cannot have if it is consistent, by Tarski's theorem. Real definability cannot be expressed in the object theory if it is consistent.

Indeed, one can turn Berry's paradox itself into another metatheoretic result. Just as Tarski showed that the truth predicate for standard arithmetic cannot be expressed in a consistent theory of arithmetic, one can show that the denotation relation ('*x* defines *y*') cannot be expressed in such a theory. This lesser-known result was established by Hilbert and Bernays.[21]

Of course, if the notion of definability for a language cannot be expressed in that language, but only in a metalanguage, Berry's paradox is not forthcoming, for exactly the same reason that the Liar paradox is not forthcoming if the notion of truth for a language cannot be expressed in that language, but only in a metalanguage. To obtain the Liar paradox, one needs all instances of the *T*-schema, $T{<}\varphi{>} \leftrightarrow \varphi$, where φ can itself contain the truth predicate. Similarly, for Berry's paradox, one needs all instances of the denotation-schema (essentially what (**) provides) where φ can itself contain the denotation predicate.

One can therefore, of course, attempt to solve Berry's paradox in exactly the same way that it was orthodox at one time to attempt to solve the Liar paradox: by appealing to a hierarchy of metalanguages. But this is not Zhong's approach. And just as well. Since Kripke's 'Outline of a Theory of Truth' [1975], it is now generally acknowledged that such a solution is unworkable.[22]

5. In Conclusion

We see, then, that Zhong's position is not sustainable. The criterion she proposes does not distinguish in a satisfactory way between the set-theoretic and the semantic paradoxes; she does not have a uniform solution, even for the semantic paradoxes; she has provided no reason why the naïve denotation relation is 'indefinite' (other than that its definiteness leads to contradiction); and her own account of the denotation relation clearly misses the mark, even by consistent standards.[23]

Universities of Melbourne and St Andrews,
and the Graduate Center
City University of New York

Received: February 2012
Revised: March 2012

References

Hilbert, D., and P. Bernays 1939. *Grundlagen der Mathematik*, Vol. II, Berlin: Springer Verlag.

Kripke, S. 1975. Outline of Theory of Truth, *Journal of Philosophy* 72/19: 690–716.

Kripke, S. 2011. A Puzzle about Time and Thought, ch. 13 of *Philosophical Troubles*, Oxford: Oxford University Press.

[21]See Hilbert and Bernays [1939: 263–78]. On this, see Priest [1997] and [2005, ch. 8].

[22]See Kripke [1975]. See also the discussion in Priest [1987, ch. 1].

[23]Many thanks to Haixia Zhong for helpful comments on an earlier draft of this paper, which eliminated a number of misunderstandings.

Priest, G. 1987. *In Contradiction: a Study of the Transconsistent*, Dordrecht: Martinus Nijhoff; second (extended) edition, Oxford: Oxford University Press, 2006.

Priest, G. 1991. Intensional Paradoxes, *Notre Dame Journal of Formal Logic* 32/2: 193–211.

Priest, G. 1995. *Beyond the Limits of Thought*, Cambridge: Cambridge University Press; second (extended) edition, Oxford: Oxford University Press, 2002.

Priest, G. 1997. On a Paradox of Hilbert and Bernays, *Journal of Philosophical Logic* 26/1: 45–56.

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

Priest, G. 2011. Comments on Kripke's Thought Paradox, a paper given at the conference on *Philosophical Troubles*, at the Graduate Center, CUNY, September 2011.

Russell, B. 1907. On Some Difficulties in the Theory of Transfinite Numbers and Order Types, *Proceedings of the London Mathematical Society* 4/1: 29–53; reprinted as ch. 7 of *Essays in Analysis*, ed. D. Lackey, London: Allen and Unwin, 1973.

Russell, B. 1908. Mathematical Logic as Based on the Theory of Types, *American Journal of Mathematics* 30/2: 222–62; reprinted as pp. 150–82 of *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, ed. J. van Heijenoort, Cambridge, MA: Harvard University Press.

Zhong, H. 2012. Definability and the Structure of Logical Paradoxes, *Australasian Journal of Philosophy* 90/4.