

Working Papers in Philosophy:
Registers of Philosophy

Szerkesztők: Kovács Gábor
Paár Tamás

2019/3

The Limits of Philosophical Rigour

The Inevitability of Paradox in Consistent Reasoning

János V. Barcsák

Pázmány Péter Catholic University



Institute of Philosophy
Research Centre for the Humanities
Hungarian Academy of Sciences

About the author

János V. Barcsák received his MA in English and aesthetics and his PhD in English romantic poetry from Eötvös Loránd University, Budapest. He also studied literary theory at the University of Oxford. Since 1994 he has taught at the Institute of English and American Studies of Pázmány Péter Catholic University, where he is currently a senior lecturer. His research focuses mainly on literary theory and English romanticism. His most recent publications include studies on Percy Bysshe Shelley's poetry, on ethical theory, and on the connections between formal logic and literary theory.

Abstract

In his book, *The Unity of Content and Form in Philosophical Writing*, Jon Stewart identifies the main reason for the impoverishment of present-day philosophical writing in the undue efforts of philosophers to emulate scientific methods, including the attempt to reduce all philosophical arguments to a rigorous logical formulation. Does the requirement of austere logical form, however, really sanction a bland and unimpassioned style in philosophical expression? In my paper I will argue that it does not. It does not because even the most rigorous foundation of logical consistency, first-order logic, must face the issue of paradox in the form of the "Gödel sentence" (the famous undecidable proposition with which Gödel proved his incompleteness theorems). I will argue that the status of the Gödel sentence in logical theories suggests that in all consistent philosophical reasoning one must face the issue of paradox insofar as one remains committed to the truth. This, I will conclude, precludes uniformity of style and unimpassioned expression in philosophical writing as long as it remains committed both to logical consistency and to truth.

**János V. Barcsák: The Limits of Philosophical Rigour
The Inevitability of Paradox in Consistent Reasoning¹**

“paradox is the passion of thought, and the thinker without a paradox is like the lover without a passion: a paltry mediocrity”²

“apparently the truth which the poet utters can be approached only in terms of paradox”³

The quotations I have chosen as the epigraphs of my paper both point at the central significance of paradox in relation to thinking the truth. In the first one, Kierkegaard argues that the encounter with absolute truth inevitably creates paradox and thus that genuine thought, thought that commits itself to the truth, will always be paradoxical. Paradox, therefore, seems to become for Kierkegaard the mark of the genuineness, of the honesty of thought. Just as a lover without a passion is not to be trusted, so a thinker without a paradox is inevitably a sham. Similarly, Cleanth Brooks argues in the passage quoted above that paradox is the necessary form in which alone poetry can convey the truth it expresses. And form here does not simply mean stylistic arrangement: what Brooks traces in the poems he analyses is *logical* form, and he claims that what underlies all poetic expressions of the truth is formal paradox, logical self-contradiction.

In my paper I would like to make a similar claim for philosophy in general. I will try to show that all philosophical reasoning is necessarily confronted with paradox produced inside its own logical structure insofar as it remains committed both to consistency and to the truth. This, I will try to demonstrate, is a *formal* requirement of reasoning, one that derives from exigencies of formal logic. I will demonstrate, in other words, that the adherence to logical consistency and scientific precision – which Jon Stewart identified as one of the main reasons for the impoverishment of present-day philosophical writing⁴ – in fact does not sanction the kind of uniformity that its application in philosophical publications has led to. Concurring with Stewart

¹ This paper was presented at the conference “Registers of Philosophy IV.,” May 26, 2018, Budapest, organized by the Institute of Philosophy of the Hungarian Academy of Sciences.

² Søren Kierkegaard, *Philosophical Fragments: Or, A Fragment of Philosophy*, eds. and trans. Howard V. Hong and Edna H. Hong, Princeton, NJ: Princeton University Press, 1985, 37. (Translation slightly modified.)

³ Cleanth Brooks, “The Language of Paradox,” in *The Well Wrought Urn: Studies in the Structure of Poetry*. London: Methuen, 1947, 3.

⁴ Jon Stewart, *The Unity of Content and Form in Philosophical Writing. The Perils of Conformity*. London: Bloomsbury, 2013, 3.

in his conclusions but approaching the matter from a different angle, I will argue that logical consistency seriously considered will require rather than eliminate paradox.⁵

Logical Consistency and Gödel Incompleteness

According to Stewart, the main philosophical reason for the current hegemony of the genre of unimpassioned, treatise-form articles in Anglophone academic philosophical discourse is the influence of logical positivism. Rudolf Carnap and A. J. Ayer advocated an approach to philosophy in which a straightforward and entirely transparent presentation of results gained with scientific methods was the only legitimate form of philosophical expression. In their view – as Stewart explains – the important thing was the scientific result of the philosopher’s investigations; the language in which this was expressed was of secondary importance and thus it was best kept to the simplest, most austere formulation.⁶ What is more, this approach also entailed the view that – as Stewart puts it – any “ideas of philosophic merit must be expressible in treatise form with rigorous, preferably formalized arguments”.⁷ Thus the logical positivists believed that any meaningful philosophical argument – whether contemporary or written in the past – could be reduced to formalized propositions and formally controlled derivations and that their real philosophical value could thus be unambiguously assessed. This view led to the dismissal of such thinkers as Hegel or Heidegger and to a general contempt for literary style in philosophical writing.⁸

Does the criterion of logical formalizability, however, really sanction this crude rejection of stylistic variety and provide an infallible measure of philosophical merit? I want to argue that it does not.

Let us examine first of all the logical positivist assumption that all meaningful philosophical arguments can be represented in a rigorous, formalized way. This would mean that the structure of any such argument can be reduced to a set of propositions and derivations in first-order logic.⁹ Of course the transition from an argument expressed in natural language to a

⁵ In *The Unity of Content and Form*, Stewart himself analyses the philosophical uses of paradox as deployments of a literary device (ibid. 31-42). However, as will become clear from the argument that follows, I use paradox in a slightly different – though by no means incompatible – meaning.

⁶ Ibid. 7.

⁷ Ibid. 9.

⁸ Ibid. 8.

⁹ Why first-order logic is the appropriate reduction-base here would take a little too long to explain and it is not unambiguous either. I will, therefore, simply assume this for now.

logical representation is never unproblematic – and logical positivists were quite aware of this, too – their assumptions, however, imply that at least in principle such a reduction is possible.

If, however, we accept this, then any rigorous philosophical argument needs to face the inevitability of paradox, for we know from Kurt Gödel’s famous meta-logical investigations that first-order logical theories are subject to incompleteness. This means that in any such theory Gödel-type undecidable propositions can be formulated. A Gödel-type undecidable proposition (let us call this for short the Gödel sentence) is essentially like semantic paradoxes in natural languages: its truth-value is undecidable. Perhaps the simplest natural-language formulation of a semantic paradox is the sentence (which is a version of the liar-paradox): ‘This statement is false.’¹⁰ It is not difficult to see that if this statement is true, then it is false, since it states precisely its own falsehood and if it is false, then it is true, since by stating its own falsehood it thus makes a true statement. The sentence, therefore, leads us to formal contradiction. What Kurt Gödel achieved in his epochal 1931 paper “On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems I”¹¹ was that he proved that such statements will crop up even in the most rigorous logical theories and that thus all such rigorously formulated theories are incomplete in the sense that meaningful statements can be formulated in them about which the theory is unable to decide – for the formal reasons outlined above – whether they are true or false.¹²

Two Gödelian Arguments in Philosophy

How does this affect philosophical reasoning? Are Gödel’s meta-logical results in any way pertinent to philosophy? Apart from its significance in the philosophical issues of mathematical research (especially for the heavily philosophical problems of the foundations of mathematics), Gödel’s incompleteness theorem also sparked several arguments which were purely philosophical. Lured by the elegance, the simplicity and the broad generalizability of Gödel’s results, many philosophers blindly referred to Gödel to bolster rather specious arguments. This blind, uninformed use of Gödel has been effectively criticized in several recent

¹⁰ See on this Willard van Orman Quine, *The Ways of Paradox and Other Essays*. New York: Random House, 1966, 9.

¹¹ Kurt Gödel, *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*, trans. B. Meltzer, New York: Dover Publications, 1992 (Basic Books, 1962).

¹² Gödel’s argument is of course crudely simplified in my brief presentation. On the relation between the liar paradox and the Gödel sentence see for example *ibid.* 40, or Douglas R. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid*. London: Penguin Books, 1979, 17-18.

books.¹³ However, there are some more serious philosophical arguments, which make use of Gödel's findings in more tenable ways. In what follows, I will briefly outline two arguments that derive from careful considerations of Gödel's results, and have received significant attention among philosophers. It is in contrast to these theories that I will then outline my own approach.

The first notable philosophical theory that rested on Gödel's meta-logical results was perhaps John R. Lucas's argument about artificial intelligence. In his 1961 paper, "Minds, Machines and Gödel" Lucas argued against the theory of "strong artificial intelligence", the belief that it is possible to model the human mind in its entirety on a machine. He pointed out that all computing machines operate in a way that is analogous to the functioning of formal systems of the type that Gödel had studied.¹⁴ Consequently, Lucas argued, such a machine will never be able to make sense of the Gödel sentence: it cannot even formulate it, because if it could mechanically produce its Gödel sentence, it would produce a falsehood (since it would derive a statement that claims its own undecidability) and that would break its consistency. For a machine, therefore, the Gödel sentence is sheer absurdity, it is totally inconceivable, but the human mind can perceive its relevance and thus we can prove that there is at least one case where a machine cannot fully model human thought processes and that there is at least one idea that the human mind can deal with but that a machine will never be able to grasp.¹⁵

Lucas's argument generated heated philosophical debates and received a lot of incisive criticism. In spite of this, however, his approach was revived on more solid foundations by Roger Penrose in his books, *The Emperor's New Mind* and *Shadows of the Mind*. To our day Penrose remains a powerful exponent of the philosophical position that human consciousness can never be fully captured by computing machines and his argument relies on Gödel's results concerning incompleteness.¹⁶

Another philosophical application of Gödel that I think deserves attention is Szaniszló (Stanley) Jáci's argument about the incompleteness of physics. In his 1966 book, *The Relevance of Physics* – and in several articles afterwards – Jáci argued that no scientific theory can ever fully capture the physical reality for the simple reason that any such theory would inevitably be

¹³ E.g.: Francesco Berto, *There's Something about Gödel: The Complete Guide to the Incompleteness Theorem*. Oxford: Wiley-Blackwell, 2009, 131-145. Alan Sokal – Jean Bricmont, *Fashionable Nonsense: Postmodern Intellectuals' Abuse of Science*. New York: Picador, 1998, 176-181. Torkel Franzén, *Gödel's Theorem: An Incomplete Guide to Its Use and Abuse*. Wellesley, MA: Peters, 2005.

¹⁴ John R. Lucas: "Minds, Machines, and Gödel," *Philosophy* 36, (1961): 113.

¹⁵ Ibid. 115.

¹⁶ Roger Penrose, *The Emperor's New Mind*. Oxford: Oxford University Press, 1989.; Roger Penrose, *Shadows of the Mind*. Oxford: Oxford University Press, 1994.; Roger Penrose: "Gödel, the Mind, and the Laws of Physics," in *Kurt Gödel and the Foundations of Mathematics*, eds. Matthias Baaz et al., Cambridge: Cambridge University Press, 2011, 339-358.

reducible to a first-order logical formulation and as such it would be susceptible to Gödel incompleteness.¹⁷ This means that the language of such a theory would contain statements which are meaningful – that is, unambiguously refer to some state of affairs in the physical universe – but about which the theory is incapable of deciding whether they are true or false, whether or not what they describe exists in the physical reality. In other words, any such theory would be unable to account for the entirety of the physical universe; there would always remain states of affairs about whose existence it could say nothing.

Unlike Lucas's hotly debated theory, Jáki's views were largely ignored by the scientific community – perhaps because they were not stated in a directly polemical context. His considerations, however, came back with a vengeance in the context of the search for the Theory of Everything (ToE), which gathered momentum in the 1980s and aimed at producing an ultimate scientific theory which would account for all observable phenomena in the physical universe. In his lectures and articles dating from this time Stephen Hawking, for example, predicted with great enthusiasm that by the turn of the millennium we would have such a theory. His high hopes, however, were quelled in 2002 by his realization – for which incidentally he did not credit Jáki – that Gödel incompleteness made it theoretically impossible to achieve this aim.¹⁸ The relevance of Gödel incompleteness to physics has recently been questioned by many scientists and philosophers of science;¹⁹ however, in the philosophy of science Gödel's results still continue to play an important role.

Incompleteness as Limitation and Excess

How is all this, however, related to philosophical writing, especially insofar as its style or its registers are concerned? For us to understand this we will need to adopt a different approach to Gödel incompleteness from the one characterizing the previous two examples.

Let us first see, however, what is common in the way these two philosophical arguments use Gödel. The first thing we must note is that in both the Gödel sentence is treated as a limitation to the systematic effort of capturing an external reality. Both approaches, therefore, start out from a sharp distinction between system and reality: between that which describes/represents and that which is present, that which merely mirrors/maps and that which *is*.

¹⁷ Stanley L. Jáki, *The Relevance of Physics*. Chicago: University of Chicago Press, 1966, 127-130.

¹⁸ Cf. Stephen Hawking's Dirac Lecture of 2002 titled "Gödel and the End of Physics," <<http://www.hawking.org.uk/godel-and-the-end-of-physics.html>> Accessed: 4 February 2019.

¹⁹ E.g.: John D. Barrow, "Gödel and Physics," in *Kurt Gödel and the Foundations of Mathematics*, eds. Matthias Baaz et al., Cambridge: Cambridge University Press, 2011, 255-276.

This dichotomy implies that the reality is prior to the merely reflective, secondary system. The latter is merely a tool with which we try to capture the former. If we fail, therefore, this will in no way affect the prior existence of what we are trying to capture. It is this priority of the real that the two approaches outlined above want to establish and maintain through their use of Gödel's incompleteness theorem. And the way they do this follows a similar logic in the two arguments, as well. The second similarity between them is, therefore, that the limitation the Gödel sentence imposes is treated in both as the mark of an excess. For Lucas and Penrose, the inevitability of the Gödel sentence attests to the non-mechanizability of thought, to the fact that "the mind can always go one better"²⁰ than a machine, while for Jáki and Hawking it secures the inexhaustibility of the empirical reality and thus ensures that scientific research can continue indefinitely. This is why – although Gödel incompleteness is a limitation on the system – all the thinkers mentioned above celebrate, rather than deplore, the fact that any consistent system is incomplete. For them the Gödel sentence guarantees the excess of the reality: the human mind is always more than what can be captured in the operation of a machine and the physical universe is always richer than what any theory can grasp.

Indeed one of the (philosophical) dangers of the scientific optimism characteristic of such approaches as Artificial Intelligence research and the search for the ToE is that, were these projects successful, they would inevitably entail the collapsing of the system/reality distinction. If the optimistic hope that the reality can be fully captured in our theories were to be fulfilled, there would be no way of telling the reality apart from its systematic description. This is what Lucas points out when he considers Alan Turing's suggestion that after reaching a critical threshold of complexity machines may cease to be predictable and acquire a form of consciousness. In the unlikely event that such a machine could be constructed, Lucas argues, it would no longer be a machine, it would be a mind indistinguishable in its performance from other minds.²¹

Although neither Jáki nor Hawking entertains this possibility, we can mount a similar argument about the potential success of the ToE project. Let us suppose that we could produce a theory that would exactly describe the physical reality in its entirety, that is, that we could find the Theory of Everything. What would the discovery of the ToE mean for physics? It would no doubt mean the end of scientific research, for if our theory indeed completely matched the reality it describes in every detail, then it would exactly and infallibly predict any outcome of any empirically observable situation. There could be no surprising results, no unexpected findings;

²⁰ Lucas, "Minds, Machines, and Gödel," 118.

²¹ Ibid. 127.

the outcome of any experiment could be decided in advance and thus there would be no need for experimentation at all. This would in effect mean that there is no difference between the reality we describe and the system in which we carry out the description. In fact, all statements about the reality would merely be reduced to statements about the inner functioning of the system. The reality, therefore, would cease to be the reality, it would collapse back upon our interpretive system, there could be no way of telling it apart from the latter.

In both these arguments, however, the danger of collapsing the system/reality opposition is defused by Gödel's theorem: since no mechanical procedure can ever make sense of the Gödel sentence, while a mind can, "the mind can always go one better" and thus remains for Lucas ultimately different from a machine; and similarly, since no physical theory can ever be complete, the physical reality will always exceed any theoretical description of it and thus, as Hawking puts it, "we will always have the challenge of new discovery."²²

These uses of the incompleteness theorem are undoubtedly in line with Gödel's own conception. Gödel was a mathematical Platonist; he believed in the non-physical but objective reality of mathematical concepts, and apparently he considered his incompleteness theorem as the foundation of such an approach.²³ In his central assessment of the philosophical significance of his incompleteness results, in the 1951 "Gibbs Lecture",²⁴ he also uses the incompleteness theorems to bolster his position that there is something real in mathematical statements. In particular, he takes issue with the conventionalist attitude (which he also terms "nominalism"), which contends that mathematical statements are true only by virtue of the conventions (the definitions and syntactical rules) that produce them and not owing to any transcendent mathematical facts that they describe. In contrast, Gödel argues that the incompleteness theorem makes this approach untenable, since the Gödel sentence will always provide a statement that the syntactic rules are incapable of controlling.²⁵ The inevitable presence of the Gödel sentence, therefore, attests to the fact that mathematics is not merely tautological, that it owes its meaning to some objective mathematical reality which exists prior to the systematic effort of grasping it.

²² Hawking. Jáki has something very similar to say in his essay, "A Late Awakening to Gödel in Physics": "Gödel's theorem remains a serious assurance to all physicists that their minds will forever be challenged by ever fresh problems." (Stanley L. Jáki, "A Late Awakening to Gödel in Physics" <<http://www.sljaki.com/texts/2004-09-a-late-awakening.pdf>> Accessed: 12 August 2019)

²³ See on this Berto, *There's Something about Gödel*, 149. See also Rebecca Goldstein, *Incompleteness: The Proof and Paradox of Kurt Gödel*, New York: Norton, 2005. Goldstein argues that mathematical Platonism was the main motivation behind all of Gödel's most important results, including the incompleteness theorems.

²⁴ Kurt Gödel, "Some basic theorems on the foundations of mathematics and their implications," in *Collected Works of Kurt Gödel*, eds. Solomon Feferman et al., vol. III, 304-323.

²⁵ Ibid. 316. In fact the incompleteness theorems provide just one line of argument that Gödel pursues in the "Gibbs Lecture". However, the way he uses them is clearly reminiscent of the Lucas/Penrose and the Jáki/Hawking arguments.

Thus for Gödel, just as for the Lucas/Penrose and the Jónsson/Hawking arguments, the Gödel sentence marks an excess which attests to the reality of what the mechanical system attempts to capture and prevents thereby the system from collapsing upon itself, from becoming a mere formal game, “an idle running of language [...] void of content.”²⁶

The Radical Immanent Approach

All these approaches, however, depend on a more or less unreflected metaphysical realism: they assume – rather than prove – the objective reality of the mind, of the empirical world, or of mathematical concepts. Gödel is of course the least guilty of this naïve assumption. At the end of his Gibbs Lecture he admits that his arguments – including his references to the incompleteness theorems – do not actually prove the Platonist (or conceptual realist) position.²⁷ He acknowledges that all he can hope to achieve with these arguments is to disprove the “nominalist” position, which maintains that mathematics is merely “syntax of language”.²⁸

In contrast, the other approaches outlined above are not so cautious in their conclusions. They derive their insights solely from syntactic considerations: they reflect on the functioning of the system and since – by Gödel’s syntactic theorem – this functioning can be shown to fail at a certain point (marked by the Gödel sentence), they draw the conclusion that this failure can only come from the excess of what the system attempts to capture (the empirical world or the human mind). Does the failure of the system alone warrant this conclusion, though? It only does if we presuppose the prior existence of the reality, and the uses of Gödel’s theorem discussed so far fail to take this fact into account.

What if, however, we do not assume the prior existence of the reality, if we start out from the hypothesis that whatever is experienced can only come to us in and through some system? This approach – which we could call a position of radical immanence – would certainly remove the naivety of the metaphysical realism entailed by the approaches discussed so far, but would it yield a feasible interpretation in their stead? And if so, how would the status of the Gödel

²⁶ Ibid. 319. That Gödel regarded the undecidability of a statement to be a mark of its reality is also attested by his definition of “analyticity”, which – he insists – does not mean that statements are true merely owing to our definitions, but because of the nature (we could say *the reality*) of the concepts occurring in them: “This concept of analytic,” he explains, “is so far from being ‘void of content’ that it is perfectly possible that an analytic proposition might be undecidable [...]. For our knowledge of the world of concepts may be as limited as that of [the] world of things” (Ibid. 321).

²⁷ Ibid. 321-322.

²⁸ See on this also Kurt Gödel, “Is Mathematics Syntax of Language?” in *Collected Works of Kurt Gödel*, eds. Solomon Feferman et al., vol. III, 334-356.

sentence change in this interpretation? These are the two questions I will explore in the remaining part of this paper.

As to the coherence or feasibility of the radical immanent attitude, we must first of all see that the radicalism involved in this approach strongly depends on the kind of reality whose presupposition we reject. The reality that is presupposed in the three theories above, for example, is quite different in each case. It is clearly not the same to claim that mathematical objects, such as the natural numbers (1, 2, 3, ...) are real beyond our theories of arithmetic, or that physical objects (such as planets or electrons) are real beyond our theories, or that the mind is more than just a machine. Similarly, the reverse hypothesis – the radical immanent assumption that the reality has no prior existence, or at least we cannot know that it does, since everything we know comes to us in and through the context of some system – yields rather different results in the three cases. It is, for example, certainly not quite as radical to claim that natural numbers only exist inside our theories of arithmetic as to contend that we observe planets or electrons to behave the way we do only because our theories allow no other perceptions. However, in none of these cases is it impossible (or incoherent) to start out from the latter hypothesis.²⁹

It is enough just to consider the success in continental philosophy of what is sometimes referred to as “the linguistic turn”, that is, the assumption that everything we can perceive of or say about “the reality” (whatever this may be) is always already inscribed in some context which functions as a language, directed by its own autonomous syntactic rules which have nothing to do with “the reality”. This approach – sometimes also referred to as post-structuralist – is perhaps best summed up by Jacques Derrida’s famous claim that “there is nothing outside of the text”³⁰ and has certainly proved to be a fruitful and inspiring approach to many philosophers working in the continental tradition. It has led to many insightful arguments, which – in spite of the detractors of post-structuralism – are certainly no less coherent than alternative depictions of the phenomena examined. What is more, it is not just in the continental tradition that the feasibility of the radical immanent position can be demonstrated. In analytic philosophy Hilary Putnam’s “internal realism”, for example, can be seen as an expression of a similar stance and, although Putnam gave up “internal realism” from 1990, his arguments put forward in the one and

²⁹ Let me remark at this point that this radical immanent approach clearly does not contradict the uses of Gödel incompleteness above, since – at least in the case of the Lucas/Penrose and the Jáci/Hawking arguments – the inexhaustibility of the mind and that of empirical reality is merely conveyed to us by the purely intra-systematic device of the Gödel sentence. Rather than the reality imposing an external limitation on our systematic thought processes, therefore, we merely infer the excess of the real from an internal anomaly within the functioning of the system.

³⁰ Jacques Derrida, *Of Grammatology*, Trans. Gayatri Chakravorty Spivak, Baltimore: Johns Hopkins University Press, 1997, 158.

a half decades preceding that date clearly attest to the feasibility and coherence of such an approach.³¹

Such a radical immanent position, however, has its own problems, too. For if we assume that what can be accessed is always just some system and never the reality, then the question of reference inevitably arises: how can we make our system refer to anything other than itself, to anything “real”? This is not a problem that can easily be dismissed in post-structuralist fashion by claiming that the system in fact *creates*, rather than describes, the reality we encounter, for this begs the question – and serious thinkers in the post-structuralist tradition are quite aware of this³² – of how we can make a distinction between different discourses, different linguistic creations of the reality. Are these all equally valid? Is there no qualitative difference between alternative accounts of the “reality”? If we admit – as I think we must – that there are better and worse arguments, superficial and profound visions, irresponsible and thoughtful discourses, then we have implicitly smuggled in a reliance on an independent reality, for what else could be the measure of these qualitative differences if not an autonomous external reality, which these discourses make better or worse reference to, which they approximate more or less closely?

The radical immanent approach, therefore, is even more vulnerable to the problems arising when a complete matching between system and reality is achieved. We saw how the utopian aim of discovering the ToE or of constructing a machine capable of self-consciousness entailed the philosophical problem of the reduction of the system/reality opposition. In spite of the difficulties that this reduction involves, however, we can still argue that in a realist account – that is, if we posit the prior existence of the reality – the indistinguishability of the system from the reality can still be made sense of. A complete isomorphism between the theory and what it describes can still function as a perfect mapping between two distinct realms whose separate existence we have presupposed – even though we can no longer account for it. As soon as we waive the realist assumption and adopt a position of radical immanence, however, the consequences of a perfect system and the resulting reduction of the system/reality opposition become truly disastrous. The indistinguishability of the system from the reality inevitably entails

³¹ The two central early statements of Putnam’s internal realism are Hilary Putnam, “Models and Reality,” *The Journal of Symbolic Logic*, 45:3 (Sep. 1980): 464-482; Hilary Putnam, *Reason, Truth and History*, Cambridge: Cambridge University Press, 1981. Putnam arrives at his position from a substantially different direction than post-structuralism (which Putnam would simply label “cultural relativism”). He considers the consequences of the Löwenheim-Skolem Theorem and concludes that language is incapable of uniquely fixing its referents and that therefore no reference to the reality can ever be transparent. Thus, he contends, the reality will always appear to us relative to our conceptual schemes and never in itself, which, he adds however, does not prevent the reality appearing in our conceptual schemes from being real.

³² Derrida for example raises the issue in his “Structure, Sign and Play” (Jacques Derrida, “Structure, Sign and Play in the Discourse of the Human Sciences,” in *Writing and Difference*, trans. Alan Bass, London: Routledge, 1978, 351-370.) and in his many discussions of “empiricism”.

a complete loss of contact with the real. Since our map, the system, is entirely autonomous and self-sufficient and since it is indistinguishable from what it maps, the reality, the very existence of the latter is called into question. If the true propositions derived in the theory are true “no matter what the facts are,”³³ just by virtue of the correct application of syntactic rules, then there is just no way of accessing any “facts”, or even of ascertaining that they *are* “out there”. We can just never be sure of what – if anything – our true propositions are true of. The system closes upon itself and allows no way out from its claustrophobic mesh.

Just as in the case of the theories above, however, Gödel’s incompleteness theorem can be invoked to salvage this situation. More precisely, even from a position of radical immanence we can rely on the Gödel sentence to mark an excess.³⁴ This we can do, moreover, without recourse to a presupposition of the reality, since the Gödel sentence can be treated as purely syntactic and remains available under the strictest mechanical control in the formal system. With the help of the Gödel sentence, in other words, we can maintain the radicalism of the immanent approach and still demonstrate the possibility of a point of excess. This excess, however, will be substantially different from that invoked in the Lucas/Penrose and the Jáki/Hawking arguments and even from that which Gödel himself attributes to it: it will not be an excess of the reality, but rather that of the system.

Indeed if the Gödel sentence were to mark the excess of the reality, then this would violate the basic assumption of the radical immanent approach, which stipulates that all that we have access to is just the system, never the reality. The Gödel sentence, however, is entirely within the system in the sense that nothing but the axioms and syntactic rules of a first-order logical theory are necessary for its constitution; and yet it is excessive insofar as it frustrates the most fundamental function of any formal system, its ability to distinguish truth from falsehood.³⁵ As we could see, inside the formal system the Gödel sentence is undecidable, it can neither be proved, nor disproved without contradiction. As such, it is the one point in the system that does not gain its meaning from its truth or falsehood, that is, from a fact in the reality which it purportedly refers to. This need not mean, however, that it does not have a meaning or referent;

³³ Gödel, “Some basic theorems,” 319.

³⁴ For an insightful analysis of possible connections between the Gödel sentence and Derrida’s key concepts see Paul M. Livingston, *The Politics of Logic. Badiou, Wittgenstein, and the Consequences of Formalism*. New York: Routledge, 2012, 122 ff. see also János V. Barcsák, “Formalization, Politics, Creativity,” in *Intertextuality, Intersubjectivity, and Narrative Identity*, ed. Péter Gaál-Szabó, Cambridge: Cambridge Scholars Publishing, 2017, 5-20.

³⁵ To be more precise, it frustrates the system’s ability to distinguish theorems from non-theorems, that is, those sentences written in the formalism of the system that can be mechanically derived from the axioms from those that cannot. This theorem-proving ability of a formal system is what makes it useful for mathematicians who turn to the formal system to make sure that their proofs are correct and their statements are really true. What the Gödel sentence in fact attests to is thus a gap between theoremhood and truth.

on the contrary, it is precisely by its ability not to refer to anything “out there” that it can come to acquire a unique reference: it comes to refer to the autonomy of the system. It indicates that not everything in the system is determined by what the system purportedly refers to, that the system is something more, something other than the reality it describes.

What does all this mean, however, from a radical immanent point of view? I have just pointed out that the fundamental assumptions of this view question the very existence of the reality; how can then the Gödel sentence – of which we can surely know that it does not refer to the reality – still restore the system/reality opposition and thus salvage the formidable philosophical problems inherent in the radical immanent approach? We saw that these formidable problems primarily arise from the way the radical immanent approach questions the very existence of the reality, which leads to its inability to account for the obvious differences in the quality of various discourses, for the fact that some arguments are more convincing, more profound, more valid than others. These problems, however, can be handled if we rely on the insight that the complicated involvement of the Gödel sentence in any consistent theory provides. For from a radical immanent position this involvement attests to the fact that the system is autonomous and thus different from what it purportedly describes. This, however, also implies that the system is not everything, that since it is distinct, there must be something that it is distinct *from*. Thus in a radical immanent approach the Gödel sentence in fact provides an intra-systematic token of transcendence.

What this transcendent thing from which the system is distinct actually *is* is of course impossible to say in a radical immanent approach, but one thing is clear: alone of all statements in the system the Gödel sentence can be seen as formulating the claim that there is something more than just the system. No other statement in the system is capable of making this claim, since in their case the referent is always taken for granted and they are therefore unable to claim explicitly – although they all presuppose this – that there is reference, that there is truth, that there exists something other than the system, that which they are entirely the products of. The Gödel sentence, by contrast, precisely by making reference to nothing else but the autonomy of the system, can be seen as asserting that the system is not just about itself, that it *can* make reference to something other, that it *can* lay claim to truth.

Moreover, since all other statements can be fully accounted for from the mechanical operation of the system, their claim to truth – as we have seen above – can easily be void of content. Since their truth or falsehood is predetermined by intra-systematic, syntactic processes, their truth claim may well be completely detached from the “facts”. The Gödel sentence, by contrast, directly frustrates the regular functioning of the system and with this it asserts that truth

and falsehood may be more than just what the system derives. It is in this sense the only intra-systematic guarantee that it is at all possible for the system to reach out to the reality and thus to state the truth.

It is important to note that this functioning of the Gödel sentence does not violate the basic assumption of the radical immanent approach: the Gödel sentence does not provide access to an actual transcendence, it merely points at an immanent pattern or structure which must inevitably shape any truth claim that the system is capable of making. It is in this way, too, that the Gödel sentence makes the radical immanent approach capable of accounting for the different qualities of discourse that we witness in philosophical arguments. An argument will be more or less true, more or less convincing, more or less profound in proportion as it engages with the inevitability of the Gödel sentence to found its own structure. To translate this claim into a less formal context, we can therefore claim that insofar as a philosophical argument remains committed both to consistency (representability in a first-order logical theory) and to truth (correct reference to something other than the descriptive system), it must – for purely formal, syntactical reasons – engage in one way or another with paradox (that is, with the Gödel sentence of its own logical construction). And this is true not only for grand scale philosophical arguments about the inexhaustibility of the physical universe or the non-mechanizability of the human mind. It is equally true for any philosophical discussion on the sole condition that it remains committed both to consistency and to truth. Without the paradox a philosophical argument is merely autotelic play, the mechanical functioning of a self-sufficient and self-absorbed system. As such it will remain merely repetitive even if it is about a seemingly new topic and will be just another product of the publishing industry that Jon Stewart deplors in present-day philosophical discourse. We can, therefore, paraphrase the Cleanth Brooks quotation that we started out from and state that the truth the philosopher utters can only be approached in terms of paradox.

This inevitability of paradox in consistent reasoning, however, is not just a fact that a thinker can simply put up with: it necessarily causes a traumatic experience. This is so because the thinker's first commitment – in most cases at least – is a commitment to consistency. The philosopher sets out to make order, to clarify, to see clearly and arrange all the facts of the case consistently. In other words, the philosopher seeks to eliminate all ambiguity and, above all, all forms of paradox. When this commitment to consistency is confronted with the formal exigencies the commitment to truth imposes, that is, when one realizes that truth can only be spoken in terms of paradox, one must react passionately. No true philosopher committed – as they must be – both to truth and to logical consistency can stay unimpassioned. We can, therefore, paraphrase Kierkegaard to state that just as a lover without a passion is not a lover at

all, a philosopher without the recognition of inevitable paradox and a passionate response to it does not think. Looked at in this way, therefore, the formal requirements of logical consistency do not authorize an unimpassioned, bland, uniform philosophical style; they rather demand passionate personal involvement and a brave espousal of paradox.

