Winner of Chadwick Prize 2010

Befriending the Liar Paradox

Mike Arnautov

This essay does not aspire to the status of a scholarly paper. Its purpose is rather to recall and record a modest adventure in the world of ideas, which came my way a few months ago. Philosophy can be, arguably, characterised as thinking about thinking (about some subject) and at one end of the spectrum such meta-thought can concern itself with a detailed examination of a narrowly delimited subject, while at the other it may range widely, offering a bird's-eye view of the terrain of ideas and being more concerned with the scope of the overall view than with its precision in detail. The train of thought, or perhaps more precisely the cascade of thought, which is described below, falls into the second category. My purpose in presenting it is to share my pleasure and excitement in finding unexpected (to me at at least) connections between apparently unrelated areas of thought. My secondary aim is to show that thinking about logic need not be a dry, arid affair and to this end I recount my mental journey in its quirky detail, instead of following the more traditional approach of presenting the end-result supported by post-hoc justifications.

My starting point was the last year's Chadwick Prize winning essay [1], which argued with ruthless efficiency that the Liar Paradox (essentially “this statement is false”) does not belong to the company of valid propositions. This is a time-honoured philosophical pursuit, and I found myself feeling sorry for the poor, much abused Liar Paradox. Generations of logicians and philosophers have done their best to squash it by any and all means available, to sweep it under the philosophical rug, to declare it a non-proposition, a persona non-grata, to invent rules of logic which would one way or another rule it out of order. One of the giants of logic even went so far as to decree that no (formal) language may contain its own truth predicate![1]

Clearly the Liar Paradox is the pariah of logic, to be suppressed, exorcised, cast into the outer darkness. And yet, reading that essay, I could not help recalling Jung's dictum that what is thrown out of the metaphorical door, will come back disruptively through the metaphorical window[2]. And indeed the Liar Paradox in its many guises is notorious for causing trouble. As is well known, Frege was discomfited by Russell's version of the Paradox phrased in the set-theoretical framework[3]. Hilbert's great formalist programme was demolished by Gödel's recasting of the Liar Paradox in a most ingenious way to prove the essential incompleteness of arithmetic [5]. And Tarski's relegation of the truth predicate of a language to the metalanguage level was also intended to preclude the vicious self-negation inherent in any fixed-point, self-referential propositions.

I have little doubt as to what Jung himself would have advised in the circumstances: befriend the Liar Paradox, do not chase it away – it may guide you to places you never dreamed of. Clearly a ridiculous thought. The Liar Paradox is a self-contradiction, a vicious form of self reference. Since self-reference is bound to feature somehow in the workings of the human mind with its effortless self-referential capability, surely it is right to seek ways to exclude the self-destructive form of self-reference embodied in all variants of the Liar Paradox.[4]

The matter would have probably rested there, except for one of those coincidences, which occasionally open doors in blank conceptual walls. As it happens, almost at the same time I came across this quote from a lecture by Sir Michael Atiyah (an eminent mathematician and a former president of the Royal Society):

> Geometry is … about space... If I look out at the audience in this room I can see a lot, in one single second or microsecond I can take in a vast amount of information... . Algebra, on the other hand... is concerned essentially with time. Whatever kind of algebra you are doing, a sequence of operations is

1 Alfred Tarski asserted that , in general, one cannot define a truth predicate "if the order of the metalanguage is at most equal to that of the language itself" [2] (as quoted by http://plato.stanford.edu/entries/tarski).

2 I appreciate that this is not a majority view, but I find Jungian and post-Jungian thought to be instrumentally very helpful, without it being necessary to underpin it with some metaphysical ontology. An “as if” attitude is quite adequate in guiding one's intuition and feeling.

3 As can be seen from the last-minute addition of an appendix to the 2nd volume of “The Basic Laws of Arithmetic” [3], in which Frege discusses the implications of Russell's Paradox.

4 This is not to contradict Quine's careful analysis (in “The Ways of Paradox” [4]) of different types of self-negation. From the bird's-eye perspective of this essay, such differences – however valid – are immaterial..

> performed one after the other, and "one after the other" means that you have got to have time. In a static universe you cannot imagine algebra, but geometry is essentially static.[5]

The assertion of algebra being intimately connected with time may seem surprising, until one realises that what is meant is not physical time measured by some real-world clock, but logical time – an ordering of sequences of successive operations. This is a kind of time programmers know well when they write program statements such as x = y + 1, which is not an equation to be solved, but an assignment to be performed. Here y *first* has some value, which is *then* increased by one, and *then* the value of x is set to this new value.

The Atiyah quote made me realise that if viewed algebraically, the Liar Paradox is not a paradox at all! Let the algebraic clock tick: if the self-referential negation is assumed to be true (tick), then it is false (tock), therefore it is true (tick) and hence it is false (tock)... At no point of algebraic time is the Liar Paradox both true *and* false, it only turns into a paradox if we attempt to squash it into a timeless single value. In fact, it is... it is... that old friend of programmers – a flip-flop![6]

At this point my thoughts attempted to go in three separate directions at once. One strand was a memory of seeing a TV interview with James Hillman (a major figure in post-Jungian depth psychology). The salient point was Hillman explaining his dislike of airports – it was all those sliding doors he did not like. Doors should open properly, pivoting on a hinge, to show both sides of the barrier of an entrance-way. Not for nothing did Romans make the two-faced God Janus to be the god of doorways. A proper door opens, and closes, hinged on a fixed point. Hillman was indicating this action by turning his hand in front of him and became visibly fascinated by this movement, observing his turning hand and murmuring: "Opening... Closing... Opening... Closing..." to the complete bemusement of the interviewer. While this memory was not illuminating either logically or philosophically, such connections matter in that they set the tone of our intuition. The flip-flop action of Hillman's hand indicated a doorway – a way through. What is more, it was presided over by a god who these days is perceived as two-faced and hence untrustworthy, and yet who used to be thought wise in being able to look both forward and back at the same time.

The second strand was all about the way 20C science was invigorated when it finally caught up with Poincaré's pioneering work of 1891[7] and started paying attention to iterative equations, thereby opening our eyes to the world of fractals and of deterministic chaos. The excitement of this period was well caught by Tom Stoppard in his play "Arcadia" [8]:

> Valentine: "[…] People were talking about the end of physics. Relativity and quantum looked as if they were going to clean out the whole problem between them. A theory of everything. But they only explained the very big and the very small. […] We are better at predicting events at the edge of the galaxy or inside the nucleus of an atom than whether it'll rain on auntie's garden party three Sundays from now. Because the problem turns out to be different. We can't even predict the next drip from a dripping tap when it gets irregular. Each drip sets up the conditions for the next, the smallest variation blows the prediction apart […] A door like this has cracked open five or six times since we got up on our hind legs. It's the best possible time to be alive, when almost everything you thought you knew is wrong."

The third strand of my thoughts took me back to my early teens and to the birthday present of an electrical experimentation kit, complete with a battery. I clearly recall the thrill of putting together an electric bell and seeing how it worked. A loosely hinged metal arm closes the bell's electric circuit, which causes an electromagnet to generate a magnetic field. This field makes the metal arm move and strike the bell, but the movement also breaks the electrical circuit, the magnetic field disappears and the arm falls back, thereby closing the circuit again. The circuit is closed (tick), therefore it is opened (tock), therefore it is closed (tick) and therefore it is open again (tock)... The arm oscillates, and yes, that's what it essentially is – an oscillator. The Liar Paradox is a logical

5 Sir Michael Atiyah's 2000 Toronto lecture, as quoted by John Derbyshire in his "Unknown Quantity" [6].

6 Most programmers would have a flip-flop in their programs somewhere, somehow, sooner or later. In many programming languages the numerical value of zero is also interpreted as the logical value of **false**, while the value of one (or in fact any non-zero value) is in a logical context interpreted as meaning **true**. The simplest flip-flop is the assignment x = 1 – x. If x is zero, it becomes one, if it is one it becomes zero. "If I am true, I am false, if I am false, I am true." For example, when right-justifying text in a non-proportional font, padding spaces are inserted alternately from from the left or from the right in successive lines of text, in order to minimise the chance of vertical "rivers". A flip-flop will be typically used to govern the left/right choice for each successive line.

7 In the early 1890s Poincaré used iterative methods to find deterministic chaos lurking within Newtonian mechanics [7]. However, science was not ready for his discovery and the whole subject had to be rediscovered independently about 70 years later.

oscillator! The arm oscillates, and the bell rings. Is it perhaps a doorbell? Who is there?

Looking back at it, I find it interesting that all three strands in their different ways involved the image of a doorway. However, what I was conscious of at the time was the inevitable voice of scepticism: does this line of thought actually lead anywhere? If not, then all this excitement does not rate even as a tiniest footnote in the long history of the Liar's Paradox. What is the big deal? If we were talking about electronics, that would be different. Practically the whole of the electronics industry is built on the foundation of the humble oscillator. Could there be such a thing as a logical circuit and could it be made to do anything of interest? Can one build such circuits beyond the obvious elaborations of the Liar Paradox (e.g. "A: statement B is true; B :statement A is false")?

There is of course a disappointingly obvious answer: logical circuits are the basis of computer industry and in fact all programming can be viewed as implementing carefully structured systems (often vast in scale) of interlocking logical statements. This is obvious from the identification of the Liar paradox with a flip-flop programming construct. Yet this answer does not feel quite right.. Flip-flops and similar software components apart, computer programs are not generally built in a way which would make them analogous to a system of propositions through which truth values may flow. Could anything interesting follow from examining such proposition systems in their own right?

If that thought sounds familiar, there is a reason for it. In response to Tarski's ruling (intended to block the Liar Paradox) that a language may not contain its own truth predicate, Kripke came up with an alternative in his "Outline of a Theory of Truth" [9]. His suggestion was to let the truth predicate initially range only over *some* particular paradox-free subset of propositions of a given language, and let logical operations propagate truth of propositions outwards from this chosen subset, until a stable point is reached. While in one way this was clearly an approach in the spirit of algebraic time, it was still motivated by the static, geometric view. The aim was to reach a stable point, beyond which all change would cease. Time was merely a technical device for getting to the timeless stability. But can it be in fact guaranteed that a stable point **would** be reached? In general this seems doubtful. Could one experiment with truth values propagating through a large, arbitrary network of propositions? On reflection, that sounded *very* familiar. Didn't Stuart Kauffman do something of the sort?

And yes, indeed he did, coming to this thought from a completely different direction. Kauffman experimented with large Boolean networks, envisaging each node of the network as a lightbulb which is either lit ("true") or not lit ("false"), the state of each bulb at the time **t+1** being a logical function of the state at time **t** of some other bulbs. Typically in such a network a lot of bulbs are either permanently on or off , but there is also a number of groups of interconnected bulbs which oscillate with different periodicity, so that the group returns to the same on/off pattern after some specific number of time steps.

That in itself does not sound too illuminating and I regret to say that reading about these experiments some years ago I simply shrugged them off. What I missed at the time was that the interesting aspect of such networks was not the behaviour of any single one of them, but the statistical features of the whole class of randomly connected networks. These properties include, for example, effects remarkably reminiscent of phase transition in physics, as one varies probabilistic parameters governing random networks. Even more interestingly, for a given set of such probabilistic parameters it turns out to be possible to predict the likely number of separate groups which do not settle into a steady on-off state, but continue to oscillate.

This is not the place to recount Kauffman's results and speculations he based on those results – they can be found in his books, of which "At Home in the Universe" [10] is probably the most approachable one. However, it is worth noting some of the implications Kauffman suggests. These range from a suggestion that we have been missing a self-organising principle in nature (what he calls "order for free", which makes emergence of life as good as inevitable), to a proposed connection between the number of different cell types in a given species and the number of genes in that species' DNA (on the assumption that networks of gene activations are governed by similar stochastic laws to those of networks of light bulbs).

Kauffman of course has his own research agenda which is not primarily about propagation of truth through systems of propositions. Connectivist AI research appears to be also encroaching on this area, but again, the primary aims of these efforts are elsewhere. Does anybody study more generally the way truth values propagate in algebraic time through systems of logical propositions? I am unaware of any such work, though that may be simply a reflection of my ignorance. Perhaps there is nothing to be found in this subject (call it "logodynamics"?) other than the statistical correlations observed by Kauffman. One possible line of enquiry could take its cue from physics and investigate effects of small perturbations on the stability of Kauffman's islands of activity. But such considerations belong to the class of "thinking about" rather than of "thinking about thinking about" and thus lie outside the scope of this essay.

Returning to a philosophical level, I find it remarkable that such rich vistas of ideas open up if we look at the Liar Paradox from a less conventional angle. Perhaps logicians and philosophers missed a trick here. In their eagerness to flatten the Paradox into a timeless, and hence clearly absurd conjunction of **true** and **false**, did they not miss an opportunity to steal a march on programmers and complexity theorists, in exploring the fascinating lands to which the Liar Paradox can lead us, if properly befriended?

References

1. Christian Michel: “The Liar Paradox Outside-In”, The Philosophical Society Review 2009, pp 47-51.
2. Alfred Tarski: "Der Wahrheitsbegriff in den formalisierten Sprachen". 1936, Studia Philosophica, 1: pp 261–405.
3. Gottlob Frege: “ Grundgesetze der Arithmetik”, 1903, Jena: Verlag Hermann Pohle, Band II.
4. W.V.Quine: “The Ways of Paradox and other essays”, Harvard University Press 1976, pp 1-18.
5. Kurt Gödel: "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I", 1931, Monatshefte für Mathematik und Physik, 38: pp173–198.
6. John Derbyshire: “Unknown Quantity: a Real and Imagined History of Algebra”, Atlantic Books 2008.
7. Henri Poincaré: “Les méthodes nouvelles de la mécanique céleste”, Paris : Gauthier-Villars, 1892 vol. I , 1893 vol. II, 1899 vol. III
8. Tom Stoppard: “Arcadia”, Faber and Faber 1993, pp 47-48.
9. Saul Kripke,: "Outline of a Theory of Truth". 1975, Journal of Philosophy, 72: pp 690–716.
10. Stuart Kauffman: “At Home in the Universe”, Oxford University Press 1996.