

**Paradoxes  
and Their Resolutions**

*A Thematic Compilation*

**Avi Sion, Ph.D.**

© COPYRIGHT AVI SION, 1990-2017.

PROTECTED BY INTERNATIONAL AND PAN-AMERICAN COPYRIGHT  
CONVENTIONS.

ALL RIGHTS RESERVED.

NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY MANNER  
WHATSOEVER,  
OR STORED IN A RETRIEVAL SYSTEM OR TRANSMITTED,  
WITHOUT EXPRESS PERMISSION OF THE AUTHOR-PUBLISHER,  
EXCEPT IN CASE OF BRIEF QUOTATIONS WITH DUE ACKNOWLEDGMENT.

Proudly self-published by Avi Sion, 2017.  
Geneva, Switzerland.

[www.TheLogician.net](http://www.TheLogician.net)  
[avi-sion@thelogician.net](mailto:avi-sion@thelogician.net)

### **Library Cataloguing Information:**

Sion, Avi

**Paradoxes and their Resolutions**

*A Thematic Compilation*

No index, no bibliography

## **Abstract**

*Paradoxes and their Resolutions* is a ‘thematic compilation’ by Avi Sion. It collects in one volume the essays that he has written in the past (over a period of some 27 years) on this subject. It comprises expositions and resolutions of many (though not all) ancient and modern paradoxes, including: the Protagoras-Euathlus paradox (Athens, 5<sup>th</sup> Cent. BCE), the Liar paradox and the Sorites paradox (both attributed to Eubulides of Miletus, 4<sup>th</sup> Cent. BCE), Russell’s paradox (UK, 1901) and its derivatives the Barber paradox and the Master Catalogue paradox (also by Russell), Grelling’s paradox (Germany, 1908), Hempel’s paradox of confirmation (USA, 1940s), and Goodman’s paradox of prediction (USA, 1955). This volume also presents and comments on some of the antinomic discourse found in some Buddhist texts (namely, in Nagarjuna, India, 2<sup>nd</sup> Cent. CE; and in the Diamond Sutra, date unknown, but probably in an early century CE).

**A is A (and not-A is not-A);  
nothing is both A and not-A;  
nothing is neither A nor not-A.**

*(Aristotle's three laws of thought.)*

## Contents

Foreword.....	8
1. The vanity of the tetralemma.....	13
1. Phenomena are positive.....	13
2. There are no negative phenomena.....	14
3. A misinterpreted experiment.....	16
4. Defining negation.....	18
2. Clarifying contradiction.....	21
1. Dialectic .....	21
2. Contradiction.....	23
3. Consistency is natural .....	29
3. Clarifying negation.....	32
1. Negation in adduction .....	32
2. Positive and negative phenomena .....	35
3. Negation is secondary .....	37
4. Negation is an intention .....	41
4. Clarifying paradox.....	44
1. Internal inconsistency.....	44
2. The Stolen Concept Fallacy .....	46
3. Systematization .....	49
4. Properties.....	53
5. The Liar paradox (early).....	59
1. Double paradox .....	59
2. The Liar paradox .....	61
3. More on the Liar paradox.....	63
4. The utility of paradoxes .....	65
6. The Liar paradox (redux).....	70
1. First approach.....	70
2. Second approach .....	72
3. Third approach .....	75
4. Fourth approach .....	77
5. Fifth approach .....	79
6. Sixth approach.....	81

7.	The Russell paradox (early) .....	84
1.	Self-membership .....	84
2.	The Russell paradox .....	86
3.	Impermutability .....	88
4.	The Barber paradox .....	92
5.	The Master Catalogue paradox.....	94
6.	Grelling's paradox .....	98
8.	The Russell paradox (redux).....	103
1.	Elements of class logic .....	103
2.	An apparent double paradox.....	109
3.	A bit of the history.....	121
4.	A bit of self-criticism.....	128
9.	More on the Russell paradox .....	131
1.	My resolution of the Russell paradox.....	131
2.	Why Russell's resolutions are inadequate ..	134
3.	Why Rescher's resolution is inadequate.....	138
4.	Other incarnations of the paradox .....	140
5.	About the Barber paradox.....	142
10.	Hempel's paradox of confirmation .....	144
1.	Traditional analysis.....	144
2.	Novel analysis.....	146
3.	The problem and its solution .....	152
11.	Goodman's paradox of prediction.....	160
1.	The alleged problem .....	160
2.	The logical solution .....	161
3.	The concept of 'grue' .....	164
4.	How far up the scale? .....	167
12.	The Sorites paradox.....	171
1.	What's a heap?.....	171
2.	The use of vague terms .....	172
3.	Reasoning with vague terms.....	175
4.	Making up fake paradoxes.....	178

13.	Protagoras vs. Euathlus.....	182
1.	An ancient paradox .....	182
2.	First resolution .....	184
3.	Second resolution .....	186
4.	Inadequate resolutions.....	190
14.	Buddhist antinomic discourse.....	195
1.	The tetralemma.....	195
2.	Neither real nor unreal .....	200
15.	More Buddhist antinomic discourse .....	207
1.	The ‘I have no thesis’ thesis.....	207
2.	Calling what is not a spade a spade.....	225
	Main References .....	230

## Foreword

*Paradoxes and their Resolutions* is a ‘thematic compilation’. It collects in one volume the essays that I have written in the past (over a period of some 27 years) on this subject.

It comprises expositions and resolutions of many, though far from all, ancient and modern paradoxes, including: the Protagoras-Euathlus paradox (Athens, 5<sup>th</sup> Cent. BCE), the Liar paradox and the Sorites paradox (both attributed to Eubulides of Miletus, 4<sup>th</sup> Cent. BCE), Russell’s paradox (UK, 1901) and its derivatives the Barber paradox and the Master Catalogue paradox (also by Russell), Grelling’s paradox (Germany, 1908), Hempel’s paradox of confirmation (USA, 1940s), and Goodman’s paradox of prediction (USA, 1955).

I also here present and comment on some of the antinomic discourse found in some Buddhist texts (namely, in Nagarjuna, India, 2<sup>nd</sup> Cent. CE; and in the Diamond Sutra, date unknown, but probably in an early century CE).

Despite its title, note well, the present book is not intended as an exhaustive study; there are many paradoxes it does not mention or treat.<sup>1</sup>

---

<sup>1</sup> In Wikipedia, [https://en.wikipedia.org/wiki/List\\_of\\_paradoxes](https://en.wikipedia.org/wiki/List_of_paradoxes), there is an interesting ‘list of paradoxes’, which groups the paradoxes under various headings and has links to entries on individual paradoxes. There are also articles on paradoxes in the online Stanford Encyclopedia of Philosophy. A book on this subject worth studying is Nicholas Rescher’s *Paradoxes: Their Roots, Range, and Resolution* (Chicago, Ill.: Open Court, 2001).



By ‘paradox’ is here, of course, meant *apparent double paradox*. A paradoxical proposition has the self-contradictory hypothetical form ‘if P, then not-P’, or ‘if not-P, then P’, where P is any sort of proposition. Such a proposition, taken alone, is not antinomic, i.e. in breach of the laws of thought, simply because there is a formal way out of it: ‘if P, then not-P’ implies the categorical proposition ‘not-P’; and ‘if not-P, then P’ implies the categorical proposition ‘P’. A *single* paradox, then, constitutes logically legitimate discourse. Very different is a *double* paradox, i.e. a claim that *both* ‘if P, then not-P’ and ‘if not-P, then P’ are simultaneously true for a given instance of P. Such a claim is diametrically opposed to the law of non-contradiction, since it concludes that both the categorical propositions ‘not-P’ and ‘P’ are simultaneously true for a given instance of P.

Clearly, resolving (double) paradoxes is essential to logic, ontology and epistemology, since to accept any such antinomy would put human knowledge, and indeed the cognitive faculties that make it possible, in grave doubt. Skeptics relish paradoxes, because they maliciously wish to invalidate human knowledge and the human mind. Defenders of human reason are therefore obligated to confront every such challenge, and neutralize it convincingly. More positively, paradoxes are great opportunities to learn something new about the way we think or what we believe, and to discover and correct our errors. If we did not encounter the paradox, we might remain unaware of our errors; the paradox opens the door to our correcting them.

‘Resolution’ of paradox consists in showing that the apparent double paradox is in fact, for some specific reason, only apparent; i.e. it is illusory, not real. To find the resolution, one generally needs to examine the underlying or surrounding discourse very carefully, and uncover where in it one made a mistake. One may have relied on

some overly vague notion or fallen into equivocation or made some unjustified assumption or whatever. Resolution of paradox is generally not a mechanical process, but requires considerable perspicacity and reflection. The main paradoxes are far from easy to dissolve; each one requires due consideration and special treatment. Often, after they are discovered by some logician or philosopher, they remain unresolved for a long time.

As regards the ordering of the present collection, I would like to make the following clarification. In my past thematic compilations (namely, *The Laws of Thought*, *The Self*, *Ethics*, and *Theology*), I have generally opted for a chronological presentation. However, in the present case, the ordering has been determined by didactic as well as chronological considerations.

- I have placed as the opening chapter an essay, “The vanity of the tetralemma,” drawn from my *A Fortiori Logic* (2013). The next two chapters, “Clarifying contradiction” and “Clarifying negation,” are drawn from my book *Ruminations* (2005). These first three essays serve to, at the outset, focus attention on and reaffirm the laws of thought.
- Next, I insert, as introduction to the subject of paradoxes, the essay “Paradoxes,” drawn from my earliest work, *Future Logic* (1990).
- This is followed by some early thoughts about paradoxes in general and the Liar paradox in particular, in the essay “The Liar paradox (early),” drawn from *Future Logic* and *Ruminations*. At this point, I insert my more advanced analysis of this paradox in “The Liar paradox (redux),” drawn from *A Fortiori Logic*.
- Next, in the essay “The Russell paradox (early),” also drawn from *Future Logic* and *Ruminations*, I present some early thoughts on the Russell paradox, as well as on certain derivatives of it, namely the barber paradox,

the book catalogue paradox and Grelling's paradox. After that, I insert my more advanced analysis of this paradox in "The Russell paradox (redux)," drawn from *A Fortiori Logic*.

- Thereafter, I have put the essays "Hempel's paradox of confirmation" and "Goodman's paradox of prediction," both drawn from my *Logical and Spiritual Reflections* (2008-9). Then the essays "The Sorites paradox is contrived" and "Protagoras vs. Euathlus resolved," both drawn from a book I have yet to complete and publish (working title: "Topics in Logic, Philosophy, and Spirituality," partly published online in 2017).
- Finally, the present volume comprises two essays, "Buddhist antinomic discourse" and "More Buddhist antinomic discourse," drawn respectively from my *Buddhist Illogic* (2002) and from *Ruminations and Logical and Spiritual Reflections*.

As for the relative importance of the paradoxes here presented and dealt with, the following may be said.

To my mind, the liar paradox and the Russell paradox are especially important, in view of their potential impact on logic theory; and I particularly recommend my most recent reflections concerning them (in the essays labeled "redux"). Although I had treated these two paradoxes in earlier writings (here labeled "early"), I managed in 2013, when I was in the last stages of writing my book *A Fortiori Logic*<sup>2</sup>, to go much more deeply into them. I was at the time at the top of my form, intellectually at my most experienced and mature, and could therefore see things a lot more clearly than ever before.

---

<sup>2</sup> In fact, these essays have nothing to do with a fortiori argument; but I happened to have written them when that book was close to finished, and I wanted to publish them as soon as possible, so I parked them in an appendix there.

In the redux essay on the liar paradox, I show that this paradox is a many-headed hydra, which cannot be explicated only with reference to self-reference, as commonly done, but involves a variety of problems that all need to be addressed. In the redux essay on the Russell paradox, I show that this class-logic conundrum is not due, as customarily assumed, to some difficulty with the idea of *non-self-membership*, but is due, on the contrary, to the impossibility of *self-membership*; this is a new and definitive resolution of the paradox, and I detail its many implications.

The modern paradoxes of Hempel and Goodman are also of significance to logic theory, in that they raise doubts in relation to induction; but their resolutions are relatively easy. The ancient sorites paradox is also logically interesting, in that it shows the importance of intellectual and verbal precision in discourse; but as a paradox it is easily resolved. The earlier Protagoras-Euathlus paradox does not have great logical importance, except perhaps to teach us that when endeavoring to resolve paradoxes we should not get caught up in logically irrelevant issues (in this case, legal ones).

As regards Buddhist antinomic discourse, dealing with it is important at this time in history, because Buddhism is fashionable in some circles, and some of its doctrines are there received in too dogmatic a spirit.

If in the future I write additional essays on paradoxes, I will hopefully include them in expanded editions of the present volume.

## 1. The vanity of the tetralemma

*From A Fortiori Logic, Appendix 7.3.*

The most radical assault on reason consists in trying to put in doubt the laws of thought, for these are indeed the foundations of all rational discourse. First, the law of identity is denied by saying that things are never quite what they seem to be, or that what they are is closer to grey than black and white. This is, of course, an absurd remark, in that for itself it lays claim to utter certainty and clarity. Then, the laws of non-contradiction and of the excluded middle<sup>1</sup> are denied by saying that things may both be and not-be, or neither be nor not-be. This is the ‘tetralemma’, the fourfold logic which is favored in Indian and Chinese philosophies, in religious mysticism, and which is increasingly referred to among some ‘scientists’. To grasp the vanity of the tetralemma, it is necessary to understand the nature of negation and the role of negation as one of the foundations of human logic.

### 1. Phenomena are positive

The first thing to understand is that everything we experience is *positive phenomenon*. Everything we perceive through our senses, or remember or imagine in our minds, or even cognize through ‘intuition’ – all that has to have some sort of content to be at all perceived. Each sense organ is a window to a distinct type of positive

---

<sup>1</sup> I inadvertently left out the words “and of the excluded middle” in the original editions of *A Fortiori Logic*, although it is clear from the rest of the sentence that I intended them.

phenomenon. We see the blue sky above, we hear birds sing, we smell the fresh air, we taste a fruit, we feel the earth's texture and warmth, etc. Similarly, the images and sounds in our heads, whether they come from memory or are produced by imagination, are positive phenomena; and even the objects of intuition must have some content that we can cognize. Secondly, we must realize that many positive phenomena may appear together in space at a given moment. This is true for each phenomenal type. Thus, the blue sky may fill only part of our field of vision, being bounded by green trees and grey buildings; we may at once hear the sounds of birds and cars; and so on. Thirdly, many positive phenomena may at any given time share the space perceived by us. Thus, superimposed on visual phenomena like the sky may be other types of phenomena: the sound of birds in the trees, the smell of traffic in the streets, the feelings in our own body, and so on. We may even hallucinate, seeming to project objects of mental perception onto physical space. For example, the image of one's eyeglasses may persist for a while after their removal. Fourthly, each positive phenomenon, whatever its type, varies in time, more or less quickly. Thus, the blue sky may turn red or dark, the sounds of birds or traffic may increase or decrease or even stop for a while, and so forth.

## **2. There are no negative phenomena**

In order to express all these perceptual possibilities – differences in space and in time and in other respects, we need a concept of negation, or more precisely an act of negating. Without 'negation', we cannot make sense of the world in a rational manner – it is the very beginning of logical ordering of our experience. Thus, in a given visual field, where (say) blue sky and trees appear, to be able to say 'the sky ends here, where the trees begin' we need the idea of 'negation' – i.e. that on one side of some boundary

sky is apparent and on the other side it is not, whereas on the first side of it trees are not apparent and on the other they are. Likewise, with regard to time, to be able to describe change, e.g. from blue sky to pink sky, we need the idea of ‘negation’ – i.e. that earlier on this part of the sky was blue and not pink, and later on it was pink and not blue. Again, we need the idea of ‘negation’ to express differences in other respects – e.g. to say that ‘the sounds of birds singing seem to emanate from the trees, rather than from buildings’. Thus, negation is one of the very first tools of logic, coming into play already at the level of sorting of experiences.

Moreover, negation continues to have a central role when we begin to deal with abstractions. Conceptual knowledge, which consists of terms and propositions based directly or indirectly on perceptual phenomena, relies for a start on our ability to cognize similarities between objects of perception: ‘this seems to resemble that somewhat’ – so we mentally project the idea of this and that ‘having something in common’, an abstract (i.e. non-phenomenal, not perceived by any means) common property, which we might choose to assign a name to. However, to take this conceptual process further, we must be able to negate – i.e. to say that ‘certain things other than this and that do not have the abstract common property which this and that seem to have’, or to say that ‘this and that do not have everything in common’. That is, we must be able to say not only that one thing resembles another in some way, but also that these or other things do not resemble each other in that way or in another way. Thus, negation is essential for making sense of information also at the conceptual level of consciousness.

Now, what is negation? To answer this question, we first need to realize that *there are no negative phenomena in the realm of experience*. Everything we perceive is positive phenomenon – because if it was not we obviously would

have nothing to perceive. We can only ‘perceive’ a negative state of affairs by first mentally defining some positive state of affairs that we should look for, and then look for it; if having looked for it assiduously we fail to find it, we then conclude inductively that it is ‘absent’, i.e. ‘not present’. Thus, positive phenomena come before negative ones, and not after. Existence logically precedes non-existence. Negative phenomena are ‘phenomena’ only metaphorically, by analogy to positive phenomena – in truth, negative phenomena are not: they do not exist. ‘Negation’ is not a concept in the sense of an abstraction from many particular experiences having a certain property in common. Negation is a tool of the thinking observer, as above described. It is an act, an intention of his.

### 3. A misinterpreted experiment

To illustrate how confused some people – even some scientists – are with regard to negation, I offer you the following example drawn from Richard Dawkins’ *The Greatest Show on Earth: The Evidence for Evolution*<sup>2</sup>. He describes an experiment by Daniel J. Simons, in which some people are asked to watch a brief video and observe how many times a certain event takes place in it; but at the end they are asked another question entirely, viz. whether they noticed the presence of a man dressed up as a gorilla in the course of the movie, and most of them admit they did not<sup>3</sup>. According to Dawkins, we may infer from this experiment how “eye witness testimony, ‘actual observation’, ‘a datum of experience’ – all are, or at least can be, hopelessly unreliable.”

---

<sup>2</sup> Pp. 13-14.

<sup>3</sup> The video can be seen at:  
[www.theinvisiblegorilla.com/videos.html](http://www.theinvisiblegorilla.com/videos.html).



But this is a wrong inference from the data at hand, because he confuses positive and negative experience. The people who watched the video were too busy looking for what they had been asked to observe to notice the gorilla. Later, when the video was shown them a second time, they did indeed spot the gorilla. There is no reason to expect us to actually experience everything which is presented to our senses. Our sensory experiences are always, necessarily, selective. The validity of sense-perception as such is not put in doubt by the limited scope of particular sense-perceptions. The proof is that it is *through further sense-perception* that we discover what we missed before. Non-perception of something does not constitute misperception, but merely incomplete perception. 'I did not see X' does not deductively imply 'I saw the absence of X', even though repetition of the former tends to inductively imply the latter.

A negative 'phenomenon' is not like a positive phenomenon, something that can directly be perceived or intuited. A negation is of necessity the product of indirect cognition, i.e. of an inductive (specifically, adductive) process. We mentally hypothesize that such and such a positive phenomenon is absent, and then test and confirm this hypothesis by repeatedly searching-for and not-finding the positive phenomenon<sup>4</sup>. If we were to at any time indeed find the positive phenomenon, the hypothesis of negation would immediately be rejected; for the reliability of a negation is far below that of a positive experience. We would not even formulate the negation, if we already had in the past or present perceived the positive phenomenon. And if we did formulate the negation, we would naturally

---

<sup>4</sup> Not-finding is the non-occurrence of the positive act of finding. Objectively, note well, not-finding is itself a negative phenomenon, and not a positive one. But subjectively, something positive may occur within us – perhaps a sense of disappointment or continued relief. See more on this topic in my *Ruminations*, chapter 9.

retract our claim if we later came across the positive phenomenon. Therefore, the content of negative phenomena is necessarily always hypothetical, i.e. tentative to some degree; it is never firm and sure as with (experienced) positive phenomena.

Negative assertions, like positive assertions, can be right or wrong. If one looked diligently for a positive phenomenon and did not find it, then one can logically claim its negation. Such claim is necessarily inductive – it is valid only so long as the positive phenomenon is actively sought and not found. The moment the positive phenomenon is observed, the negation ceases to be justified. If one did *not* look for the positive phenomenon, or did not look with all due diligence, perhaps because of some distraction (as in the example cited above), then of course the claim of negation is open to doubt; certainly, it is inductively weak, and one is very likely to be proved wrong through some later observation.

#### **4. Defining negation**

How, then, is negation to be defined? We could well say that negation is defined by the laws of non-contradiction and of the excluded middle. That is, with regard to any term ‘X’ and its negation ‘not-X’, the relation between them is by definition the disjunction “Either X or not-X” – which is here taken to mean that these terms (X and not-X) cannot be both true and cannot be both false, i.e. they are exclusive and exhaustive. What do I mean here by ‘definition’? – is that an arbitrary act? No – it is ‘pointing to’ something evident; it is ‘intentional’. Here, it points to the instrument of rational discourse which we need, so as to order experience and produce consistent conceptual derivatives from it. The needed instrument has to be thus and thus constructed; another construct than this one would not do the job we need it to do for us. That is, the only conceivable

way for us to logically order our knowledge is by means of negation defined by means of the laws of non-contradiction and of the excluded middle. Without this tool, analysis of experience is impossible.

Suppose now that someone comes along and nevertheless objects to the preceding assertion. Well, he says, how do you *know* that the dilemma “either X or not-X” is true? You just arbitrarily *defined* things that way, but it does not mean it is a fact! Could we not equally well claim the tetralemma “Either X or not-X *or both or neither*” to be true? The reply to that objection is very simple. Suppose I accept this criticism and agree to the tetralemma. Now, let me divide this fourfold disjunction, putting on the one side the single alternative ‘X’ and on the other side the triple alternative ‘not-X or both or neither’. I now again have a dilemma, viz. “either ‘X’ or ‘not-X or both or neither’.” Let me next define a new concept of negation on this basis, such that we get a disjunction of two alternatives instead of four. Let us call the complex second alternative ‘not-X or both or neither’ of this disjunction ‘NOT-X’ and call it ‘the super-negation of X’.

Thus, now, the objector and I agree that the disjunction “either X or NOT-X” is exclusive and exhaustive. We agree, presumably, that this new dilemma cannot in turn be opposed by a tetralemma of the form “Either X or NOT-X or both or neither” – for if such opposition was tried again it could surely be countered by another division and redefinition. We cannot reasonably repeat that process ad infinitum; to do so would be tantamount to blocking all rational thought forever. Having thus blocked all avenues to thought, the objector could not claim to have a better thought, or any thought at all. There is thus no profit in further objection. Thus, the tetralemma is merely a tease, for we were quite able to parry the blow. Having come to an agreement that the new disjunction “Either X or NOT-X” is logically unassailable, we must admit that the original

disjunction “Either X or not-X” was logically sound from the first. For I can tell you that what I meant by not-X, or the ‘negation of X’, was from the beginning what is now intended by NOT-X, or the ‘super-negation of X’!

I was never interested in a relative, weak negation, but from the start sought an absolute, strong negation. For such utter negation, and nothing less radical, is the tool we all need to order experience and develop conceptual knowledge in a consistent and effective manner. In other words, whatever weaker version of negation someone tries to invent<sup>5</sup>, we can still propose a strong version such that both the laws of non-contradiction and of the excluded middle are applicable without doubt to it. If such negation did not exist, it would have to be invented. No one can destroy it by denying it or diluting it. Those who try to are merely sophists who do not understand the source, nature and function of negation in human discourse. They think it is a matter of symbolic manipulation, and fail to realize that its role in human discourse is far more fundamental and complex than that. Negation is the indispensable instrument for any attempt at knowledge beyond pure perception.

---

<sup>5</sup> There are people who say that the law of non-contradiction is logically necessary, but the law of the excluded middle is not. Clearly, this claim can be refuted in the same way. If they claim the three alternatives “Either X or not-X or ‘neither X nor not-X’” – we can again split the disjunction into two, with on one side “X” and on the other side “not-X or ‘neither X nor not-X’” – and then proceed as we did for the tetralemma. The same can be done if anyone accepts the law of the excluded middle but rejects the law of non-contradiction. All such attempts are fallacious nonsense.

## 2. Clarifying contradiction

*From Ruminations 1, 5, & 9.*

### 1. Dialectic

The three “Laws of Thought” may be briefly explicated as follows:

- (i) *Thesis*: there are certain appearances; appearances appear.
- (ii) *Antithesis*: there are incompatibilities between certain of these appearances; in such cases, one or both of them must be false.
- (iii) *Synthesis*: some remaining appearances must be true; find out which!

We can in this perspective consider *dialectic* as a fundamental form of thought, through which knowledge is made to progress on and on. It is not a mere detail, an occasional thought-process, but a driving force, an engine, of thought.

The laws are not mere information, but calls to cognitive action. They enjoin proactive and curative cognitive measures, to ensure (as much as possible at any given time) continued verification, consistency and completeness.

(i) The law of identity tells us to seek out the facts and sort them out as well as we can. The purpose of this law is to instill in people a healthy respect for facts, in the course of observation and judgment. It is essentially a call to honesty, and submission to the verdict of truth. People often think, or act as if they think, that ignoring or denying unpleasant

facts or arguments will make them ‘go away’ – the law of identity says ‘no, they will not disappear, you must take them into consideration’.

Some people think that it is impossible for us to ignore that “A is A”. Far from it! All of us often do so – as when we refuse to look at or admit the evidence or a logical demonstration; when we avoid reality or evade it having glimpsed it; when we lie to ourselves or to others; and so forth. If the law of identity were always obeyed by us, there would be no need to formulate it. Logic states the obvious, because it is often shunned.

(ii) When the law of non-contradiction says to us “you cannot at once both affirm and deny a proposition”, it is also telling us that if we ever in the course of discourse encounter a situation where a proposition seems both true (for some reason) and false (for other reasons), ***we have to go back upstream in our discourse and find out where we went wrong in the course of it<sup>8</sup>, and we have to effect an appropriate correction such as to eliminate the difficulty.***

We are not just saying: “ah, there is a contradiction”, and leaving it at that, nonplussed. No, we are impelled to seek a solution to the problem, i.e. to resolve the contradiction. We are inferring that there must be something wrong in our earlier thinking that led us to this conundrum, some error of observation or reasoning that requires treatment. So long as this situation is tolerated, and we cannot pinpoint the source of error, the credibility of all related knowledge is proportionately diminished. Consistency must be restored as soon as possible, or we risk putting all subsequent knowledge in doubt.

(iii) Similarly, the law of the excluded middle does not just inform us that “no proposition can be claimed neither true nor false”. This law insists that if we find ourselves in such

---

<sup>8</sup>

“Check your premises”, Ayn Rand would say.

a situation, and it is indeed the case that both a proposition and its exact negation both seem false, we cannot let the matter rest or hope to find some compromise position – we have to eventually, as soon as possible, find good reason to opt for one side or the other. There is no logically acceptable middle ground, no avenue of escape.

These action implications inherent in the laws of thought may also be characterized as dialectical thinking. In this perspective, the “thesis” is our knowledge (or opinion) as it happens to be at a given time; the “antithesis” is the discovery of a logical flaw in that thesis, which causes us to have doubts about it and seek its review; and finally, the “synthesis” is the corrections we make in our premises, so as to resolve the difficulty encountered and obtain a less problematic new state of knowledge.

## 2. Contradiction

Many people misunderstand what we logicians mean by ‘**contradiction**’. The contradictory of a term ‘A’ is its negation, ‘not A’, which refers to anything and everything in the universe other than A, i.e. wherever precisely A is absent in the world. The relation of contradiction between A and not-A is mutual, reversible, perfectly symmetrical.

The presence of something (A) excludes its absence (i.e. not A) in that very same thing, and vice versa, if all coordinates of space and time are identical. However, this does not exclude the logical possibility that the same thing may be partly A and partly not A. Thus, the law of thought ‘either A or not A’ can also be stated more quantitatively as “either ‘all A’ or ‘all not A’ or ‘part A and part not A’”.

Some people appeal to this possibility of three alternatives as an argument *against* the laws of thought! But that is a misunderstanding – or worse, deliberate sophistry.

If something, e.g. 'B', implies but is not implied by not-A, it (i.e. B) is as 'incompatible' with A as not-A is, but it is not contradictory to A: it is merely *contrary* to A. The contradictory not-A of A differs from A's contraries in that *the absence* of not-A implies A, whereas in the case of mere contraries like B (or B1 or B2... etc.) this added logical relation of 'exhaustiveness' does not apply.

When contradictories are placed in a disjunction, 'either A or not-A', the disjunction involved signifies both mutual exclusion ('or', meaning 'not together') and exhaustiveness ('either', meaning 'and there is no other alternative'). It intends: if 'A', then not 'not-A'; and if not 'A', then 'not-A'.

On the other hand, any number of contraries can be placed in a disjunction: 'A or B or B1 or B2... etc.', so that the presence of any disjunct implies the absence of all the others; but such disjunction is not exhaustive, unless we specify that the list of contraries in it is complete. If that list *is* indeed complete, then the negation of all but one of the disjuncts implies the affirmation of the remaining one. Thus, 'not-A' can be equated to the exhaustive disjunction of all things in the world 'contrary to A'.

Something *different* from A, e.g. 'C', is not necessarily contradictory or even contrary to A. **The mere fact of difference does not imply incompatibility.** Different things (like A and C) may be compatible, i.e. capable of coexistence in the same thing, at the same time and place. 'Difference' simply signifies that we are *able to distinguish* between the things concerned: i.e. they are not one and the same when they appear before our consciousness. 'Similar' things may be the same in appearance, but not one (e.g. two instances of the same kind); or they may be one (i.e. parts of a single whole), yet not the same.

Thus, for example, the logical relation between the colors black and white depends on how precisely we focus on



them. They are different, since distinguishable. Since they may coexist on different parts of the same surface, they are broadly compatible. However, as such or *per se*, they are contrary; that is to say: if I perceive a surface or part of surface as totally white, and you perceive *the very same* place and time as totally black, our claims are incompatible<sup>9</sup>. This irreconcilability is not a contradiction, however, because it is possible for a surface to be neither black nor white.

The expression ‘**contradiction in terms**’ refers to a compound term composed of incompatible elements, such as ‘A and not A’ or ‘A and B (where B is contrary to A)’. Such a mixed-up term may be said to be paradoxical, as well as internally inconsistent, since it implies that contradiction is possible, so that the laws of thought are denied by it, and then (by generalization, if you like) ‘anything goes’ including denial of the ‘A and not A’ conjunction.

For example, the term “illusory reality” is a contradiction in terms. On the other hand, note, terms like ‘an inhuman human’ or ‘an anti-Semitic Jew’ are not strictly speaking contradictions in terms; they refer to natural possibilities of conjunction, only the terminology used makes them superficially seem contradictory (i.e. there are people who behave inhumanly, or Jews that hate their own people).

The proposition ‘A is not A’ (or ‘some thing that is A is also not A’), being self-contradictory, implies ‘A is A’, its contradictory form. This statement should be explicitly acknowledged, though obvious, because it correlates two important concepts, viz. ‘internal inconsistency’ and ‘the logic of paradoxes’.

---

<sup>9</sup> Our disagreement is not terminological, note. We have in the past agreed as to what experiences ‘black’ and ‘white’ correspond to; here, we suddenly diverge.

The statement ‘A is not A’ is logically impossible, because it both affirms and denies the same thing. Therefore, the opposite statement is true. That statement, i.e. ‘A is A’, is logically necessary, because *even its contradictory* ‘A is not A’ implies it.

Whoever claims ‘A is not A’ is admitting ‘A is A’ – *ipse dixit*, he himself said it! Whereas, whoever claims ‘A is A’ is consistent with himself.

**Self-contradiction** consists of three items:

- (i) The proposition in question, call it P.
- (ii) The admission that it is *an assertoric statement*, i.e. one that affirms or denies something.
- (iii) The admission that all assertoric statements *involve claims* to consciousness, to knowledge, to truth, etc.

Thus, given P (e.g. “reality is unknowable”), admit that P implies “this is an assertion” – but all assertions imply some knowledge of reality – *therefore*, P implies non-P. There is a process from P to its negation, which Logic demands we acknowledge. That demand cannot be refused without committing the very same self-contradiction. This is not a circular or ad infinitum proof, but an appeal to honesty, without which no dialogue is possible.

That all assertoric propositions assert is an aspect of the Law of Identity. The Law of Non-contradiction may be discerned in the argument: All assertions assert something; P is an assertion; therefore, P asserts; whence, if P denies asserting, P implies non-P. The Law of the Excluded Middle is also implicit here, in the awareness that we have no choice but to firmly disown P.

**Disguised contradictions.** Contradictions appear in discourse in many guises. They are not always overt, but may be hidden in the fact of making a statement or in the standards of judgment used.

A claim may be paradoxical because it **inherently entails** its own contradiction, although it does not on the surface seem to be self-inconsistent. Such implication is not always formal but requires awareness of the meaning of the terms used. This form of indirect self-contradiction has been called “the Stolen Concept fallacy”<sup>10</sup>.

For instance, the skeptical claim “I know nothing” may be rejected as self-contradictory, because as soon as someone makes it – someone *who understands and intends the meaning* of the terms “I”, “know” and “nothing” – that is by itself proof absolute that the person concerned “knows” *something*, whence the original claim (of total ignorance) is shown up to be unavoidably contradictory and thus necessarily false.

Thus, in cases of this sort, the tacit implication involved is that one of the terms used (knowing nothing) implicitly includes the act in question (knowing that I know nothing), as a case in point contradictory to the explicit claim. (Rephrasing the said statement as “I do *not* know anything” does not change its underlying assumptions, needless to say.)

There are countless examples of such inherent self-contradiction. Saying “I have nothing to say” is saying something. Claiming “We have no memory” is self-contradictory, because each term in it presupposes a word, concept and background experiences remembered by the speaker – and the hearer too. An amusing common example is “I do not speak a word of English”!

Another important form of covert self-inconsistency is the use of a **double standard**. This consists in applying less stringent standards of judgment to one’s own discourse than to the discourse of one’s intellectual opponents. A lot

---

<sup>10</sup> By Ayn Rand and (I think) Nathaniel Branden.

of philosophical, and particularly political and religious, discourse resorts to such inequitable methodology.

The contradiction involved in a double standard is apparent the moment we step back and view its user's knowledge and methodology as a whole. In this wider perspective, the user of a double standard is clearly inconsistent with himself, even if his discourse viewed piecemeal may superficially seem self-consistent.

Whole philosophies may be based on such fallacious reasoning. For instance, Phenomenalism sets as a general standard a limitation of knowledge to sensory data without allowing extrapolations from them to assumed external material objects – yet it does not criticize its own adductions using the same rigid standard.

There are two ways this fallacy may be committed: one may use relaxed standards on one's own discourse, while seemingly applying universal norms to one's opponents' discourse; or one may appear to apply universal norms to oneself, while concocting overly strict norms for them. One may *exempt oneself* from the usual logical rules, or one may make unusual logical *demands on others*.

In either case, the holder of a double standard is in conflict with logic's requirement of uniformity. An assumption of reason is that all humans are epistemologically on the same plane. Equity is an aspect of 'common sense'. Experience and logic have to be used to convince oneself and others, not sophisticated manipulation or authority.

Standards of judgment have to be fair and universal; all discourse must be equally treated. If differences are advocated, they have to be convincingly justified. The principle of equality admittedly involves generalization; but the onus of proof is on any proposed particularization of it.

An example of a double standard is the appeal to cultural relativism. One may seek to rationalize ideas or thought

processes that are contrary to ordinary reason, by claiming them to belong to a different cultural framework. Such tolerance seems on the surface friendly and open-minded, but it is proposed without full consideration of its negative human and epistemological implications.

### **3. Consistency is natural**

It is important to here reiterate the principle that consistency is natural; whereas inconsistency is exceptional.

Some modern logicians have come up with the notion of “proving consistency” – but this notion is misconceived. Consistency is the natural state of affairs in knowledge; it requires no (deductive) proof and we are incapable of providing such proof, since it would be ‘placing the cart before the horse’. The only possible ‘proof’ of consistency is that no inconsistency has been encountered. Consistency is an inductive given, which is very rarely overturned. All our knowledge may be and must be assumed consistent, unless and until there is reason to believe otherwise.

In short: harmony generally reigns unnoticed, while conflicts erupt occasionally to our surprise. One might well wonder now if this principle is itself consistent with the principle herein defended that negatives are never per se objects of cognition, but only exist by denial of the corresponding positives. Our principle that consistency is taken for granted seems to imply that we on occasion have logical insights of *in*consistency, something negative!

To resolve this issue, we must again emphasize the distinction between pure experience and the *interpretations* of experience that we, wordlessly (by mere intention) or explicitly, habitually infuse into our experiences. Generally, almost as soon as we experience something, we immediately start interpreting it, dynamically relating it to

the rest of our knowledge thus far. Every experience almost unavoidably generates in us strings of associations, explanations, etc.

The contradictions we sometimes come across in our knowledge do not concern our pure experiences (which are necessarily harmonious, since they in fact exist side by side – we might add, quite ‘happily’). *Our contradictions are necessarily contradictions between an interpretation and a pure experience, or between two interpretations.* Contradictions do not, strictly speaking, reveal difficulties in the raw data of knowledge, but merely in the hypotheses that we conceived concerning such data.

Contradictions are thus to be blamed on reason, not on experience. This does not mean that reason is necessarily faulty, but only that it is fallible. Contradictions ought not be viewed as tragic proofs of our ignorance and stupidity – but as helpful indicators that we have misinterpreted something somewhere, and that this needs reinterpretation. These indicators are precisely one of the main tools used by the faculty of reason to control the quality of beliefs. The resolution of a contradiction is just new interpretation.

How we know that two theories, or a theory and some raw data, are ‘in contradiction’ with each other is a moot question. We dismiss this query rather facetiously by referring to “logical insight”. Such insight is partly ‘experiential’, since it is based on scrutiny of the evidence and doctrines at hand. But it is clearly not entirely empirical and involves abstract factors. ‘Contradiction’ is, after all, an abstraction. I believe the answer to this question is largely given in the psychological analysis of negation.

There is an introspective sense that *conflicting intentions* are involved. Thus, the ‘logical insight’ that there is inconsistency is not essentially insight into a negative (a non-consistency), but into a positive (the intuitive experience of conflict of intentions). Although the word

inconsistency involves a negative prefix, it brings to mind something empirically positive – a felt tension between two theses or a thesis and some data.

For this reason, to say that ‘consistency is assumable, until if ever inconsistency be found’ is consistent with our claim that ‘negations are not purely empirical’. (Notice incidentally that we did not here “prove” consistency, but merely *recovered* it by clarifying the theses involved.)

The above analysis also further clarifies how the law of non-contradiction is expressed in practice. It does not sort out experiences as such, but concerns more abstract items of knowledge. To understand it fully, we must be aware of the underlying intentions. A similar analysis may be proposed to explain the law of the excluded middle.

In the latter case, we would insist that (by the law of identity) ‘things are something, what they are, whatever that happen to be’. Things cannot be said to be *neither* this *nor* the negation of this, because such characterizations are negative (and, respectively, doubly negative) – and therefore cannot constitute or be claimed as positive experience. Such situations refer to uncertainties *in the knower*, which he is called upon to eventually fill-in. They cannot be proclaimed final knowledge (as some modern sophists have tried to do), but must be considered temporary postures in the pursuit of knowledge.

### 3. Clarifying negation

*From Ruminations 9.*

#### 1. Negation in adduction

Concepts and theories are hypothetical constructs. They cannot (for the most part) be proven (definitely, once and for all), but only repeatedly confirmed by experience. This is the positive side of adduction, presenting evidence in support of rational constructs. This positive aspect is of course indispensable, for without some concrete evidence an abstraction is no more than a figment of the imagination, a wild speculation. The more evidence we adduce for it, the more reliable our concept or theory.

But, as Francis Bacon realized, the account of adduction thus far proposed does not do it justice. Just as important as the positive side of providing evidence, is the negative aspect of it, the rejection of hypotheses that make predictions conflicting with experience. As he pointed out, even if a hypothesis has numerous confirmations, it suffices for it to have *one* such wrong prediction for it to be rejected.

Stepping back, this means that the process of adduction is concerned with selection of the most probable hypothesis among two or more (already or yet to be conceived) explanations of fact. Each of them may have numerous 'positive instances' (i.e. empirical evidence that supports it); and so long as they are all still competitive, we may prefer those with the most such instances. But, the way we decisively advance in our conceptual/theoretical



knowledge is by the successive *elimination* of propositions that turn out to have ‘negative instances’ (i.e. empirical evidence against them).

Now all the above is well known and need not be elucidated further. This theory of inductive logic has proven extremely successful in modern times, constituting the foundation of the scientific method.

But upon reflection, the matter is not as simple and straightforward as it seems at first!

Consider, for example, the issue of whether or not there is water on Mars. It would seem that the proposition “There is water on Mars” is far easier to prove inductively than the contradictory proposition “There is no water on Mars”. Both propositions are hypotheses.

The positive thesis would be somewhat confirmed, if it was discovered using certain instruments from a distance that there are serious indices that water is present; the thesis would be more solidly confirmed, if a sample of Mars was brought back to Earth and found upon analysis to contain water. In either case, the presence of water on Mars would remain to some (however tiny) degree unsure, because some objection to our instrumental assumptions might later be raised or the sample brought back may later be found to have been contaminated on the way over. Nevertheless, something pretty close to certainty is conceivable in this matter.

The negative thesis, by contrast, is much more difficult to prove by experience. We can readily assume it to the extent that the positive thesis has not so far been greatly confirmed. That is, so long as we have not found evidence for the positive thesis (i.e. water on Mars), we should rather opt for the negative thesis. But the latter is only reliable to the degree that we tried and failed to confirm the former. If we earnestly searched for water every which way we could

think of, and did not find any, we can with proportionate confidence assume there is no water.

Thus, in our example, the negative thesis is actually *more difficult* to establish than the positive one. It *depends on a generalization*, a movement of thought from “Wherever and however we looked for water on Mars, *none was found*” to “*There is no* water on Mars”. However, note well, it remains conceivable that a drop of water be found one day somewhere else on Mars, centuries after we concluded there was none.

Granting this analysis, it is clear that Bacon’s razor that “What is important is the negative instance” is a bit simplistic. It assumes that a negative is as accessible as (if not, indeed, more accessible than) a positive, which is not always the case.

In practice, a negative may be inductively more remote than a positive. Granting this conclusion, the question arises – is the negative instance *ever* more empirically accessible than (or even as accessible as) the positive one? That is, *when* does Bacon’s formulation of induction actually come into play?

If we look at major historical examples of rejection of theories, our doubt may subsist. For example, Newtonian mechanics was in place for centuries, till it was put in doubt by the discovery of the constancy of the velocity of light (which gave rise to Relativity theory) and later again by the discovery of various subatomic phenomena (which gave rise to Quantum mechanics). In this example, the ‘negative instances’ were essentially ‘positive instances’ – the only thing ‘negative’ about them was just their negation of the Newtonian worldview!

Such reflections have led me to suspect that the ‘negation’ referred to by Bacon is only meant *relatively* to some selected abstraction. His razor ought not be taken as an advocacy of absolute negation. If we look at the matter

more clearly, we realize that the data used to thus negate an idea is essentially positive. A deeper consideration of the nature of negation is therefore patently called for.

## **2. Positive and negative phenomena**

People have always considered that there is a difference between a positive and a negative term. Indeed, that is why logicians have named them differently. But logicians have also found it difficult to express that difference substantially. Yet, there are significant phenomenological differences between positive and negative phenomena.

a. The concrete material and mental world is evidently composed only of positive particular phenomena, some of which we perceive (whether through the bodily senses or in our minds). These exist at least as appearances, though some turn out to seem real and others illusory. This is an obvious phenomenological, epistemological and ontological truth.

To say of phenomena that they are ‘particular’ is to express awareness that they are always limited in space and time. They have presence, but they are finite and transient, i.e. manifestly characterized by diversity and change.

We do not ordinarily experience anything concrete that stretches uniformly into infinity and eternity (though such totality of existence might well exist, and indeed mystics claim to attain consciousness of it in deep meditation, characterizing it as “the eternal present”). We do commonly consider some things as so widespread. ‘Existence’ is regarded as the substratum of all existents; ‘the universe’ refers to the sum total of all existents; and we think of ‘space-time’ as defining the extension of all existents. But only ‘existence’ may be classed as an experience (a quality found in all existents); ‘the universe’ and ‘space-time’ must be admitted as abstractions.

However, the limits of particulars are perceivable without need of negation of what lies beyond them, simply due to the variable concentration of consciousness, i.e. the direction of focus of attention. That is, though ‘pointing’ to some positive phenomenon (e.g. so as to name it) requires some negation (we mean “this, but not that”), one can notice the limits of that phenomenon independently of negation.

b. Negative phenomena (and likewise abstracts, whether positive or negative), on the other hand, do depend for their existence on a Subject/Agent – a cognizing ‘person’ (or synonymously: a self or soul or spirit) with consciousness and volition looking out for some remembered or imagined positive phenomenon and failing to perceive it (or in the case of abstracts, comparing and contrasting particulars).

Thus, negative particular phenomena (and more generally, abstracts) have a special, more ‘relative’ kind of existence. They are not as independent of the Subject as positive particular phenomena. That does not mean they are, in a Kantian sense, ‘a priori’ or ‘transcendental’, or purely ‘subjective’ – but it does mean that they are ontological potentials that are only realized in the context of (rational) cognition.

Another kind of experience is required for such realization – the self-experience of the Subject, his intuitive knowledge of his cognitions and volitions. This kind of experience, being immediate, may be positive or negative without logical difficulty. The Subject reasons inductively as follows:

I am searching for X;

*I do not* find X;

Therefore, X “*is not*” there.

*The negative conclusion may be ‘true’ or ‘false’, just like a positive perception or conclusion. It is true to the degree*

that the premises are true – i.e. that the alleged search for X was diligent (intelligent, imaginative, well-organized, attentive and thorough), and that the alleged failure to find X is not dishonest (a lie designed to fool oneself or others).

Whence it is fair to assert that, unlike some positive terms, negative terms are never based *only* on perception; they *necessarily* involve a thought-process – the previous mental projection or at least intention of the positive term they negate.

This epistemological truth does reflect an ontological truth – the truth that the ‘absences’ of phenomena lack phenomenal aspects. A ‘no’ is not a sort of ‘yes’.

Note well the logical difference between ‘**not perceiving X**’ and ‘**perceiving not X**’. We do not have direct experience of the latter, but can only indirectly claim it by way of *inductive inference* (or extrapolation) from the former. In the case of a positive, such process of reasoning is not needed – one often can and does ‘perceive X’ directly.

Suppose we draw a square of opposition for the propositions (labeling them by analogy to standard positions) – “I perceive X” (A), “I do not perceive not X” (I), “I perceive not X” (E), “I do not perceive X” (O). Here, the A form is knowable by experience, whereas the I form is knowable perhaps only by deductive implication from it. On the negative side, however, the E form is not knowable by experience, but only by inductive generalization from the O form (which is based on experience).

### 3. Negation is secondary

Negation is a pillar of both deductive and inductive logic, and requires careful analysis. We have to realize that negative terms are fundamentally distinct from positive ones, if we are to begin fathoming the nature of logic. The

following observation seems to me crucial for such an analysis:

***We can experience something positive without having first experienced (or thought about) its negation, but we cannot experience something negative without first thinking about (and therefore previously having somewhat experienced) the corresponding positive.***

a. Cognition at its simplest is perception. Our perceptions are always *of positive particulars*. The contents of our most basic cognitions are phenomenal sights, sounds, smells, tastes, and touch and other bodily sensations that seemingly arise through our sense organs interactions with matter – or mental equivalents of these phenomena that seemingly arise through memory of sensory experiences, or in imaginary re-combinations of such supposed memories.

A positive particular can be experienced directly and passively. We can just sit back, as it were, and receptively observe whatever happens to come in our field of vision or hearing, etc. This is what we do in meditation. We do not have to actively think of (remember or visualize or conceptualize) something else in order to have such a positive experience. Of course, such observation may well in practice be complicated by thoughts (preverbal or verbal) – but it is possible in some cases to have a pure experience. This must logically be admitted, if concepts are to be based on percepts.

b. In the case of *negative particulars*, the situation is radically different. A negative particular has *no* specific phenomenal content, but is *entirely* defined by the ‘absence’ of the phenomenal contents that constitute some positive particular. If I look into my material or mental surroundings, I will always see present phenomena. The absence of some phenomenon is only noticeable if we first

think of that positive phenomenon, and wonder whether it is present.

It is accurate to say that our finding it absent reflects an empirical truth or fact – but it is a fact that we simply would not notice the negative without having first thought of the positive. Negative knowledge is thus necessarily (by logical necessity) more indirect and active. It remains (at its best) perfectly grounded in experience – but such negative experience requires a rational process (whether verbal or otherwise).

To experience a negative, I must first imagine (remember or invent) a certain positive experience; then I must look out and see (or hear or whatever) whether or not this image matches my current experience; and only then (if it indeed happens not to) can I conclude to have “experienced” a negative.

Thinking about X may be considered as positioning oneself into a vantage point from which one can (in a manner of speaking) experience not-X. If one does not first place one’s attention on X, one cannot possibly experience the negation of X. One may well experience all sorts of weird and wonderful things, but not specifically not-X.

From this reflection, we may say that whereas affirmatives can be experienced, negatives are inherently rational acts (involving imagination, experience and intention). A negative necessarily involves thought: the thought of the corresponding positive (the imaginative element), the testing of its presence or absence (the experiential element) and the rational conclusion of “negation” (the intentional element).

c. The negation process may involve words, though it does not have to.

Suppose I have some momentary experience of sights, sounds, etc. and label this positive particular “X”. The *content of consciousness* on which I base the term X is a

specific set of positive phenomenal experiences, i.e. physical and/or mental percepts. Whenever I can speak of this X, I mentally *intend* an object of a certain color and shape that moves around in certain ways, emitting certain sounds, etc.

Quite different is the negation of such a simple term, “not X”. The latter is not definable by any specific percepts – it *refers to no perceptible qualities*. It cannot be identified with the positive phenomena that happen to be present in the absence of those constituting X. Thus, strictly speaking, not-X is only definable by ‘negation’ of X.

Note well, it would not be accurate to say (except *ex post facto*) that not-X refers to all experiences other than X (such as Y, Z, A, B, etc.), because when I look for X here and now and fail to find it, I am only referring to present experience within my current range and not to all possible such experiences. We would not label a situation devoid of X as “not X” *without thinking of X*; instead, we would label that situation in a positive manner (as “Y”, or “Z”, or whatever).

Thus, we can name (or wordlessly think of) something concrete “X”, *after* experiencing phenomena that constitute it; but in the case of “not-X”, we necessarily conjure the name (or a wordless thought) of it *before* we experience it.

“Not-X” is thus already a concept rather than a percept, even in cases where “X” refers to a mere percept (and all the more so when “X” itself involves some abstraction – as it usually does). The concept “not X” is hypothetically constructed first and then confirmed by the attempted and failed re-experience of X.

In short, negation – even at the most perceptual level – involves an adductive process. It is never a mere experience. A negative term never intends the simple perception of some negative thing, but consists of a



hypothesis with some perceptual confirmation. Negation is always conceptual as well as perceptual in status.

A theory cannot be refuted before it is formulated – similarly, X cannot be found absent unless we first think of X.

#### 4. Negation is an intention

Now, there is no specific phenomenal experience behind the word “not”. Negation has no special color and shape, or sound or smell or taste or feel, whether real or illusory! What then is it? I suggest the following:

Negation as such refers to a ‘mental act’ – or more precisely put, it is an act of volition (or more precisely still, of velleity) by a Subject of consciousness. Specifically, negation is *an intention*. Note that our will to negate is itself *a positive act*, even though our intention by it is to negate something else.

Negation does express an experience – the ‘failure’ to find something one has searched for. Some cognitive result is willfully pursued (perception of some positive phenomenon), but remains wanting (this experience is qualitatively a suffering of sorts, but still a positive intention, note) – whence we mentally (or more precisely, by intention) mark the thing as ‘absent’, i.e. we construct an idea of ‘negation’ of the thing sought.

Thus, negation is *not a phenomenon* (a physical or mental percept), *but something intuited* (an event of will within the cognizing Subject). ‘Intuition’ here, note well, means the self-knowledge of the Subject of consciousness and Agent of volition. This is experience of a *non-phenomenal* sort. Such self-experience is immediate: we have no distance to bridge in space or time.

When a Subject denies the presence of a material or mental phenomenon, having sought for it in experience and not

found it – the ‘denial’ consists of a special act of intention. This intention is what we call ‘negation’ or ‘rejection of a hypothesis’. It occurs in the Subject, though it is about the Object.

This intention is not however an arbitrary act. If it were, it would be purely subjective. This act (at its best) remains sufficiently dependent on perception to be judged ‘objective’. The Subject must still look and see whether X is present; if that positive experience does not follow his empirical test, he concludes the absence of X.

Indeed, an initial negation may on closer scrutiny be found erroneous, i.e. we sometimes think something is ‘not there’ and then after further research find it on the contrary ‘there’. Thus, this theory of negation should not be construed as a claim that our negating something makes it so. Negation is regulated by the principles of adduction – it is based on appearance that is credible so long as confirmed, but may later be belied.

We can *ex post facto* speak of an objective absence, but we cannot fully define ‘absence’ other than as ‘non-presence’, and the ‘non-’ herein is not a phenomenon but an intention. The ‘absence’ is indeed experienced, but it is *imperceptible* without the Subject posing the prior question ‘is X present?’

Absence, then, is not produced by the Subject, but is made perceptible by his vain search for presence. For, to repeat, not-X is not experienced as a specific content of consciousness – but as a continuing failure to experience the particular positive phenomena that define X for us.

Although we are directly only aware of apparent existents, we can inductively infer non-apparent existents from the experience that appearances come and go and may change. On this basis, we consider the categories ‘existence’ and ‘appearance’ as unequal, and the former as broader than the latter. Similarly, we inductively infer ‘objective absence’

from ‘having sought but not found’, even though we have no direct access to former but only indirect access by extrapolation from the latter. Such inference is valid, with a degree of probability proportional to our exercise of due diligence.

For these reasons, I consider the act of negation as an important key to understanding the nature and status of logic. Negation is so fundamental to reason, so crucial an epistemic fact, that it cannot be reduced to something else.

We can describe it *roughly* as an intention to ‘cross-off’ (under the influence of some reason or other) the proposed item from our mental list of existents. But this is bound to seem like a circular definition, or a repetition of same using synonyms. It is evident that *we cannot talk about negation without engaging in it*. Thus, we had better admit the act of negation as a primary concept for logical science.

Note in passing: the present theory of negation provides biology with an interesting distinction regarding rational animals.

Sentient beings without this faculty of negation can only respond to the present, whereas once this faculty appears in an organism (as it did in the human species) it can mentally go beyond the here and now. A merely sensory animal just reacts to current events, whereas a man can fear dangers and prepare for them.

Once the faculty of negation appears, the mind can start *abstracting, conceiving alternatives and hypothesizing*. Memory and imagination are required to project a proposed positive idea, but the intent to negate is also required to reject inadequate projections. Without such critical ability, our fantasies would quickly lead us into destructive situations.

## 4. Clarifying paradox

*From Future Logic 31.*

A very important field of logic is that dealing with paradox, for it provides us with a powerful tool for establishing some of the most fundamental certainties of this science. It allows us to claim for epistemology and ontology the status of true sciences, instead of mere speculative digressions. This elegant doctrine may be viewed as part of the study of axioms.

### 1. Internal inconsistency

Consider the hypothetical form ‘If P, then Q’, which is an essential part of the language of logic. It was defined as ‘P and nonQ is an impossible conjunction’.

It is axiomatic that the conjunction of any proposition P and its negation nonP is impossible; thus, a proposition P and its negation nonP cannot be both true. An obvious corollary of this, obtained by regarding nonP as the proposition under consideration instead of P, is that the conjunction of any proposition nonP and its negation not-nonP is impossible; thus, a proposition P and its negation nonP cannot be both false.

So, the Law of Identity could be formulated as, “For any proposition, ‘If P, then P’ is true, and ‘If nonP, then nonP’ is true”. The Laws of Contradiction and of the Excluded Middle could be stated: “For any proposition, ‘If P, then not-nonP’ is true (P and nonP are incompatible), and ‘If not-nonP, then P’ is true (nonP and P are exhaustive)”.

Now, consider the paradoxical propositions 'If P, then nonP' or 'If nonP, then P'. Such propositions appear at first sight to be obviously impossible, necessarily false, antinomies.

But let us inspect their meanings more closely. The former states 'P and (not not)P is impossible', which simply means 'P is impossible'. The latter states 'nonP and not P is impossible', which simply means 'nonP is impossible'. Put in this defining format, these statements no longer seem antinomic! They merely inform us that the proposition P, or nonP, as the case may be, contains an intrinsic flaw, an internal contradiction, a property of self-denial.

From this we see that there may be propositions which are logically self-destructive, and which logically support their own negations. Let us then put forward the following definitions. A proposition is *self-contradictory* if it denies itself, i.e. implies its own negation. A proposition is therefore *self-evident* if its negation is self-contradictory, i.e. if it is implied by its own negation.

Thus, the proposition 'If P, then nonP' informs us that P is self-contradictory (and so logically impossible), and that nonP is self-evident (and so logically necessary). Likewise, the proposition 'If nonP, then P' informs us that nonP is self-contradictory, and that P is self-evident.

The existence of paradoxes is not in any way indicative of a formal flaw. The *paradox*, the hypothetical proposition itself, is not antinomic. It may be true or false, like any other proposition. Granting its truth, it is its antecedent thesis which is antinomic, and false, as it denies itself; the consequent thesis is then true.

If the paradoxical proposition 'If P, then nonP' is true, then its contradictory 'If P, not-then nonP', meaning 'P is not impossible', is false; and if the latter is true, the former is false. Likewise, 'If nonP, then P' may be contradicted by 'If nonP, not-then P', meaning 'nonP is not impossible'.

The two paradoxes ‘If P, then nonP’ and ‘If nonP, then P’ are contrary to each other, since they imply the necessity of incompatibles, respectively nonP and P. Thus, although such propositions taken singly are not antinomic, double paradox, a situation where both of these paradoxical propositions are true at once, is unacceptable to logic.

In contrast to positive hypotheticals, negative hypotheticals do not have the capability of expressing paradoxes. The propositions ‘If P, not-then P’ and ‘If nonP, not-then nonP’ are *not* meaningful or logically conceivable or ever true. Note this well, such propositions are formally false. Since a form like ‘If P, not-then Q’ is defined with reference to a positive conjunction as ‘{P and nonQ} is possible’, we cannot without antinomy substitute P for Q here (to say ‘{P and nonP} is possible’), or nonP for P and Q (to say ‘{nonP and not-nonP} is possible’).

It follows that the proposition ‘if P, then nonP’ does not imply the lowercase form ‘if P, not-then P’, and the proposition ‘if nonP, then P’ does not imply the lowercase form ‘if nonP, not-then nonP’. That is, in the context of paradox, hypothetical propositions behave abnormally, and not like contingency-based forms.

This should not surprise us, since the self-contradictory is logically impossible and the self-evident is logically necessary. Since paradoxical propositions involve incontinent theses and antitheses, they are subject to the laws specific to such basis.

The implications and consistency of all this will be looked into presently.

## **2. The Stolen Concept Fallacy**

Paradoxical propositions actually occur in practice; moreover, they provide us with some highly significant results. Here are some examples:

- denial, or even doubt, of the laws of logic conceals an appeal to those very axioms, implying that the denial rather than the assertion is to be believed;
- denial of man's ability to know any reality objectively, itself constitutes a claim to knowledge of a fact of reality;
- denial of validity to man's perception, or his conceptual power, or reasoning, all such skeptical claims presuppose the utilization of and trust in the very faculties put in doubt;
- denial on principle of all generalization, necessity, or absolutes, is itself a claim to a general, necessary, and absolute, truth.
- denial of the existence of 'universals', does not itself bypass the problem of universals, since it appeals to some itself, namely, 'universals', 'do not', and 'exist'.

More details on these and other paradoxes, may be found scattered throughout the text. Thus, the uncovering of paradox is an oft-used and important logical technique. The writer Ayn Rand laid great emphasis on this method of rejecting skeptical philosophies, by showing that *they implicitly appeal to concepts which they try to explicitly deny*; she called this 'the fallacy of the Stolen Concept'.

A way to understand the workings of paradox, is to view it in the context of dilemma. A self-evident proposition P could be stated as 'Whether P is affirmed or denied, it is true'; an absolute truth is something which turns out to be true whatever our initial assumptions.

This can be written as a constructive argument whose left horn is the axiomatic proposition of P's identity with itself, and whose right horn is the paradox of nonP's self-contradiction; the minor premise is the axiom of thorough contradiction between the antecedents P and nonP; and the conclusion, the consequent P's absolute truth.

If P, then P — and — if nonP, then P  
 but either P or nonP  
 hence, P.

A destructive version can equally well be formulated, using the contrapositive form of identity, ‘If nonP, then nonP’, as left horn, with the same result.

If nonP, then nonP — and — if nonP, then P  
 but either not-nonP or nonP  
 hence, not-nonP, that is, P.

The conclusion ‘P’ here, signifies that P is logically necessary, not merely that P is true, note well; this follows from the formal necessity of the minor premise, the disjunction of P and nonP, assuming the right horn to be well established.

Another way to understand paradox is to view it in terms of knowledge contexts. Reading the paradox ‘if nonP, then P’ as ‘all contexts with nonP are contexts with P’, and the identity ‘if P, then P’ as ‘all contexts with P are contexts with P’, we can infer that ‘all contexts are with P’, meaning that P is logically necessary.

We can in similar ways deal with the paradox ‘if P, then nonP’, to obtain the conclusion ‘nonP’, or better still: P is impossible. The process of resolving a paradox, by drawing out its implicit categorical conclusions, may be called *dialectic*.

Note in passing that the abridged expression of simple dilemma, in a single proposition, now becomes more comprehensible. The compound proposition ‘If P, then {Q and nonQ}’ simply means ‘nonP’; ‘If nonP, then {Q and nonQ}’ means ‘P’; ‘If (or whether) P or nonP, then Q’ means ‘Q’; and ‘If (or whether) P or nonP, then nonQ’ means ‘nonQ’. Such propositions could also be categorized as paradoxical, even though the contradiction generated concerns another thesis.



However, remember, the above two forms should not be confused with the lesser, negative hypothetical, relations ‘Whether P or nonP, (not-then not) Q’ or ‘Whether P or nonP, (not-then not) nonQ’, respectively, which are not paradoxical, unless there are conditions under which they rise to the level of positive hypotheticals.

### 3. Systematization

Normally, we presume our information already free of self-evident or self-contradictory theses, whereas in abnormal situations, as with paradox, necessary or impossible theses are formally acceptable eventualities.

A hypothetical of the primary form ‘If P, then Q’ was defined as ‘P and nonQ are impossible together’. But there are several ways in which this situation might arise. Either (i) both the theses, P and nonQ, are individually contingent, and only their conjunction is impossible — this is the normal situation. Or (ii) the conjunction is impossible because one or the other of the theses is individually impossible, while the remaining one is individually possible, i.e. contingent or necessary; or because both are individually impossible — these situations engender paradox.

Likewise, a hypothetical of the contradictory primary form ‘If P, not-then Q’ was defined as ‘P and nonQ are possible together’. But there are several ways this situation might arise. Either (i) both the theses, P and nonQ, and also their conjunction, are all contingent — this is the normal situation. Or (ii) one or the other of them is individually not only possible but necessary, while the remaining one is individually contingent, so that their conjunction remains contingent; or both are individually necessary, so that their conjunction is also not only possible but necessary — these situations engender paradox.

These alternatives are clarified by the following tables, for these primary forms, and also for their derivatives involving one or both antitheses. The term ‘possible’ of course means ‘contingent or necessary’, it is the common ground between the two. We will here use the symbols ‘**N**’ for necessary, ‘**C**’ for contingent (meaning possible but unnecessary), and ‘**M**’ for impossible. The combinations are numbered for ease of reference. The symmetries in these tables ensure their completeness.

Table 1. Modalities of Theses and Conjunctions.

No.	Theses				Conjunctions			
	P	nonP	Q	nonQ	P and Q	P and nonQ	nonP and Q	nonP and nonQ
Normal (P,Q both contingent)								
1.	C	C	C	C	C	C	C	C
2.	C	C	C	C	M	C	C	C
3.	C	C	C	C	C	M	C	C
4.	C	C	C	C	C	C	M	C
5.	C	C	C	C	C	C	C	M
6.	C	C	C	C	C	M	M	C
7.	C	C	C	C	M	C	C	M
Abnormal (one or both of P, Q not contingent)								
8.	M	N	C	C	M	M	C	C
9.	N	M	C	C	C	C	M	M
10.	C	C	M	N	M	C	M	C
11.	C	C	N	M	C	M	C	M
12.	N	M	N	M	N	M	M	M
13.	N	M	M	N	M	N	M	M
14.	M	N	N	M	M	M	N	M
15.	M	N	M	N	M	M	M	N

The following table follows from the preceding. ‘**Yes**’ indicates that an implication and its contraposite are implicit in the form concerned, while ‘**no**’ indicates that they are excluded from it. ‘ $\rightarrow$ ’ here means implies, and ‘ $\leftarrow$ ’ means is implied by.

Table 2. Corresponding Definite Hypotheticals.

No.	Name	Implications ( $\rightarrow$ ) and Contrapositives ( $\leftarrow$ )			
		$P \rightarrow \text{non}Q$	$P \rightarrow Q$	$\text{non}P \rightarrow \text{non}Q$	$\text{non}P \rightarrow Q$
		$\text{non}P \leftarrow Q$	$\text{non}P \leftarrow \text{non}Q$	$P \leftarrow Q$	$P \leftarrow \text{non}Q$
Normal (P, Q both contingent)					
1.	Neutral	no	no	no	no
2.	Contrary	yes	no	no	no
3.	Subalternating	no	yes	no	no
4.	Subalternated	no	no	yes	no
5.	Subcontrary	no	no	no	yes
6.	Implicant	no	yes	yes	no
7.	Contradictory	yes	no	no	yes
Abnormal (one or both of P, Q not contingent)					
8.	P impossible, Q contingent	yes	yes	no	no
9.	P necessary, Q contingent	no	no	yes	yes
10.	P contingent, Q impossible	yes	no	yes	no
11.	P contingent, Q necessary	no	yes	no	yes
12.	P, Q both necessary	no	yes	yes	yes
13.	P necessary, Q impossible	yes	no	yes	yes
14.	P impossible, Q necessary	yes	yes	no	yes
15.	P, Q both impossible	yes	yes	yes	no

Normal hypothetical logic thus assumes the theses of hypotheticals always both contingent, and so limits itself to cases Nos. **1** to **7** in the above tables. However, the abnormal cases Nos. **8** to **15**, in which one or both theses are not contingent (that is, are self-evident or self-contradictory), should also be considered, to develop a complete logic of hypotheticals.

The definition of the primary positive form ‘If P, then Q’, while remaining unchanged as ‘P plus nonQ is not possible’, is now seen to more precisely comprise the following situations: Nos. **3**, **6**, **8**, **11**, **12**, **14**, or **15**, that is,

all the cases where 'P and nonQ' is impossible ('M'), or 'P implies Q' is marked 'yes'.

The definition of the primary negative form 'If P, not-then Q', while remaining unchanged as 'P plus nonQ is not impossible', is now seen to more precisely comprise the following situations: Nos. **1, 2, 4, 5, 7, 9, 10, or 13**, that is, all the cases where 'P and nonQ' is contingent (C), or 'P implies Q' is marked 'no'.

The other six hypothetical forms, involving the antitheses of P and/or Q, can likewise be given improved definitions, by reference to the above tables.

Notice the symmetries in these tables. In case No. **1**, all conjunctions are 'C' and all implications are 'no'. In cases Nos. **2-5**, one conjunction is 'M', and one implication is 'yes'. In cases **6-11**, two conjunctions are 'M', and two implications are 'yes'. In cases Nos. **12-15**, three conjunctions are 'M', and three implications are 'yes'. Note the corresponding statuses of individual theses in each case.

The process of contraposition is universally applicable to all hypotheticals, positive or negative, normal or abnormal, for it proceeds directly from the definitions. For this reason, in the above tables, each implication is firmly coupled with a contraposite. Likewise, the negation of any implication engenders the negation of its contraposite, so that the above tables also indirectly concern negative hypotheticals, note well.

We must be careful, in developing our theory of hypothetical propositions, to clearly formulate the breadth and limits of application of any process under consideration, and specify the exceptions if any to its rules. The validity or invalidity of logical processes often depends on whether we are focusing on normal or abnormal forms, though in some cases these two classes of proposition behave in the same way. If these distinctions

are not kept in mind, we can easily become guilty of formal inconsistencies.

#### **4. Properties**

Paradoxical propositions obey the laws of logic which happen to be applicable to all hypotheticals, that is, to hypotheticals of unspecified basis. But paradoxicals, being incontingency-based hypotheticals, have properties which normal hypotheticals lack, or lack properties which normal hypotheticals have. In such situations, where differences in logical properties occur, general hypothetical logic follows the weaker case.

The similarities and differences in formal behavior have already been dealt with in appropriate detail in the relevant chapters, but some are reviewed here in order to underscore the role played by paradox.

##### **a. Opposition.**

In the doctrine of opposition, we claimed that ‘If P, then Q’ and ‘If P, then nonQ’ must be contrary, because if P was true, Q and nonQ would both be true, an absurdity. However, had we placed these propositions in a destructive dilemma, as below, we would have obtained a legitimate argument:

If P, then Q — and — if P, then nonQ  
 but either nonQ or Q  
 hence nonP

Likewise, ‘If P, then Q’ and ‘If nonP, then Q’ could be fitted in a valid simple constructive dilemma, yielding Q, instead of arguing as we did that they must be contrary because their contrapositions result in the absurdity of nonQ implying nonP and P.

It follows that these contrarities are only valid conditionally, for contingency-based hypotheticals. There are exceptional circumstances in which they do not hold, namely relative to abnormal hypotheticals (including paradoxicals).

This is also independently clear from the observation of ‘yes’ marks standing parallel, in cases Nos. **8**, **14**, **15** (allowing for both ‘P implies nonQ’ and ‘P implies Q’, where P is impossible), and in cases Nos. **11**, **12**, **14** (allowing for both ‘P implies Q’ and ‘nonP implies Q’, where Q is necessary).

Similar restrictions follow automatically for the subcontrariety between ‘If P, not-then nonQ’ and ‘If P, not-then Q’, and likewise for the subalternation by the uppercase ‘If P, then Q’ of the lowercase ‘If P, not-then nonQ’ (which corresponds to obversion). These oppositions only hold true for normal hypotheticals; when dealing with abnormal hypotheticals (and therefore in general logic), we must for the sake of consistency regard the said propositions as neutral to each other.

### **b.     Eduction.**

Similarly with the derivative eductions. The primary process of contraposition is unconditional, applicable to all hypotheticals, but the other processes can be criticized in the same way as above, by forming valid simple dilemmas, using the source proposition and the denial of the proposed target, or the contraposite(s) of one or the other or both, as horns.

Alternatively, these propositions can be combined in a syllogism, yielding a paradoxical conclusion. Thus:

In the case of obversion or obverted conversion (in the former, negate contraposite of target):

If Q, then nonP (negation of target)

if P, then Q (source)

so, if P, then nonP (paradox = nonP)

In the case of conversion by negation or obverted inversion  
(in the latter, negate contrapositive of target):

If P, then Q (source)

if nonQ, then P (negation of target)

so, if nonQ, then Q (paradox = Q)

Thus, eductive processes other than contraposition are only good for contingency-based hypotheticals, and may not be imitated in the abnormal logic of paradoxes. This is made clear in the above tables, as follows.

Consider the paradigmatic form 'If P, then Q'. If we limit our attention to cases Nos. 1-7, then it occurs in only two situations, subalternating (3) or implicance (6). In these two situations, 'P implies nonQ' is uniformly 'no', so the obverse, 'If P, not-then nonQ' is true; and the contrapositive 'Q implies nonP' is also 'no', so the obverted converse, 'If Q, not-then nonP' is true; 'nonP implies Q' is uniformly 'no', so the obverted inverse 'If nonP, not-then Q' is true; and the contrapositive 'nonQ implies P' is also 'no', so the converse by negation 'If nonQ, not-then P' is true. With regard to inversion and conversion, they are not applicable, because 'nonP implies nonQ' and 'Q implies P' are 'no' in one case, but 'yes' in the other.

However, if now we expand our attention to include cases Nos. 8-15, we see that 'If P, then Q' occurs additionally if P is self-contradictory and Q is contingent (8) or P is contingent and Q is self-evident (11) or P, Q are each self-evident (12) or P is self-contradictory and Q is self-evident (14) or P, Q are each self-contradictory (15). The above-mentioned uniformities, which made the stated eductions feasible, now no longer hold. There is a mix of 'no' and 'yes' in the available alternatives which inhibits such eductions.

### c. Deduction.

With regard to syllogism, the *nonsubaltern moods*, validated by *reductio ad absurdum*, remain universally valid, since such indirect reduction is essentially contraposition, and no other eductive process was assumed. But the *subaltern moods* in all three figures, are only valid for normal hypotheticals. Since these moods presuppose subalternations for their validation, i.e. depend on direct reductions through obversion or obverted inversion, they are not valid for abnormal hypotheticals.

With regard to apodosis, the moods with a modal minor premise provide us with the entry-point into abnormal logic. As for dilemma, it is the instrument *par excellence* for unearthing paradoxes in the course of everyday reasoning. If we put *any* simple dilemma, constructive (as below) or destructive (*mutatis mutandis*), in syllogistic form, we obtain a paradoxical conclusion:

If P, then R — and — if Q, then R  
 but P and/or Q  
 hence, R

This implies the sorites:

If nonR, then nonP (contrapose left horn)  
 if nonP, then Q (minor)  
 if Q, then R (right horn)  
 hence, if nonR, then R (paradoxical conclusion = R)

Thus, paradoxical propositions are an integral part of general hypothetical logic, not some weird appendix. They highlight the essential continuity between syllogism and simple dilemma, the latter being reducible to the former.

It follows incidentally that, since (as earlier seen) apodosis may be viewed as a special, limiting case of simple



dilemma, and simple dilemma as a special, limiting case of complex dilemma — all the inferential processes relating to hypotheticals are closely related.

The paradox generated by simple dilemma of course depends for its truth on the truth of the premises. We should not hurriedly infer, from the paradox inherent in every simple dilemma, that all truths are ultimately self-evident, and all falsehoods ultimately self-contradictory. Knowledge is not a purely rational enterprise, but depends largely on empirical findings.

As already pointed out, simple dilemma yields a categorical necessity or impossibility as its conclusion, only if all its premises are themselves indubitably incontinent. Should there be tacit conditions for, or any doubt regarding the unconditionality of, the hypotheticals (the horns) and/or the disjunction (the minor premise), then the conclusion would be proportionately weakened with regard to its logical modality.

Thus, with reference to the foregoing example, *granting the horns* of the major premise: in the specific case where our minor premise is a formally given disjunction — if, say, P and Q are contradictory to each other ( $P = \text{non}Q$ ,  $Q = \text{non}P$ ) — then the R conclusion is indeed necessary. But usually, the listed alternatives P and Q are only contextually exhaustive, so that the R conclusion is only factually true.

So, although every logical necessity is self-evident, and every logical impossibility is self-contradictory, formally speaking, according to our definitions, we might be wise to say that these predications are not in practice reciprocal, and make a distinction between apodictic and factual paradox. The former is independently obvious; the latter derives from more empirical data, and therefore, though contextually trustworthy, has a bit less weight and finality.

Note lastly, the inconsistency of two 'equally cogent' simple dilemmas can now be better understood, as due to their implying contrary paradoxes.

## 5. The Liar paradox (early)

*From Future Logic 32, and Ruminations 1 & 5.*

### 1. Double paradox

We have seen that logical propositions of the form ‘if P, then nonP’ (which equals to ‘nonP’) or ‘if nonP, then P’ (which equals to ‘P’), are perfectly legal. They signify that the antecedent is self-contradictory and logically impossible, and that the consequent is self-evident and logically necessary. As propositions in themselves, they are in no way antinomic; it is one of their constituents which is absurd.

Although either of those propositions, occurring alone, is formally quite acceptable and capable of truth, they can never be both true: they are irreconcilable contraries and their conjunction is formally impossible. For if they were ever both true, then both P and nonP would be implied true.

We must therefore distinguish between *single paradox*, which has (more precisely than previously suggested) the form ‘if P, then nonP; but if nonP, not-then P; whence nonP’, or the form ‘if nonP, then P; but if P, not-then nonP; whence P’ — and *double paradox*, which has the form ‘if P, then nonP, *and* if nonP, then P’.

Single paradox is, to repeat, within the bounds of logic, whereas double paradox is beyond those bounds. The former may well be true; the latter always signifies an error of reasoning. Yet, one might interject, double paradox occurs often enough in practice! However, that does not

make it right, any more than the occurrence of other kinds of error in practice make them true.

Double paradox is made possible, as we shall see, by a hidden *misuse of concepts*. It is sophistry par excellence, in that we get the superficial illusion of a meaningful statement yielding results contrary to reason. But upon further scrutiny, we can detect that some fallacy was involved, such as ambiguity or equivocation, which means that in fact the seeming contradiction never occurred.

Logic demands that *either or both* of the hypothetical propositions which constituted the double paradox, or paradox upon paradox, *be false*. Whereas single paradox is *resolved*, by concluding the consequent categorically, without denying the antecedent-consequent connection — double paradox is *dissolved*, by showing that one or both of the single paradoxes involved are untrue, nonexistent. Note well the difference in problem solution: resolution ‘explains’ the single paradox, whereas dissolution ‘explains away’ the double paradox.

The double paradox *serves to show* that we are making a mistake of some kind; the fact that we have come to a contradiction, is our index and proof enough that we have made a wrong assumption of sorts. Our ability to intuit logical connections correctly is not put in doubt, because the initial judgment was too rushed, without pondering the terms involved. Once the concepts involved are clarified, it is the rational faculty itself which pronounces the judgment against its previous impression of connection.

It must be understood that every double paradox (as indeed every single paradox), is *teaching us something*. Such events must not be regarded as threats to reason, which put logic as a whole in doubt; but simply as lessons. They are sources of information, they reveal to us certain logical rules of concept formation, which we would otherwise not

have noticed. They show us the outer limits of linguistic propriety.

## 2. The Liar paradox

An ancient example of double paradox is the well-known ‘Liar Paradox’, discovered by Eubulides, a 4th cent. BCE Greek of the Megarian School. It goes: ‘does a man who says that he is now lying speak truly?’ The implications seem to be that if he is lying, he speaks truly, and if he is not lying, he speaks truly.

Here, the conceptual mistake underlying the difficulty is that the proposition is *defined by reference to itself*. The liar paradox is how we discover that such concepts are not allowed.

The word ‘now’ (which defines the proposition itself as its own subject) is being used with reference to something which is not yet in existence, whose seeming existence is only made possible by it. Thus, in fact, the word is empty of specific referents in the case at hand. The word ‘now’ is indeed usually meaningful, in that in other situations it has precise referents; but in this case it is used before we have anything to point to as a subject of discourse. It looks and sounds like a word, but it is no more than that.

A more modern and clearer version of this paradox is ‘this proposition is false’, because it brings out the indicative function of the word ‘now’ in the word ‘this’.

The word ‘this’ accompanies our pointings and presupposes that there is something to point to already there. It cannot create a referent for itself out of nothing. This is the useful lesson taught us by the liar paradox. We may well use the word ‘this’ to point to another word ‘this’; but not to itself. Thus, I can say to you ‘this “this”, which is in the proposition “this proposition is false”’, without

difficulty, because my ‘this’ has a referent, albeit an empty symbol; but the original ‘this’ is meaningless.

Furthermore, the implications of this version seem to be that ‘if the proposition is true, it is false, and if it is false, it is true’. However, upon closer inspection we see that the expression ‘the proposition’ or ‘it’ has a different meaning in antecedents and consequent.

If, for the sake of argument, we understand those implications as: if this proposition is false, then this proposition is true; and if this proposition is true, then this proposition is false — taking the ‘this’ in the sense of *self-reference* by every thesis — then we see that the theses do not in fact have one and the same subject, and are only presumed to be in contradiction.

They are not formally so, any more than, for any P1 and P2, ‘P1 is true’ and ‘P2 is false’ are in contradiction. The implications are not logically required, and thus the two paradoxes are dissolved. There is no self-contradiction, neither in ‘this proposition is false’ nor of course in ‘this proposition is true’; they are simply meaningless, because the indicatives they use are without reference.

Let us, alternatively, try to read these implications as: if ‘this proposition is false’ is true, then that proposition is false; and if that proposition is false, then that proposition is true’ — taking the first ‘this’ as self-reference and the ‘thats’ thereafter as all pointing us backwards to the original proposition and not to the later theses themselves. In other words, we mean: if ‘this proposition is false’ is true, then ‘this proposition is false’ is false, and if ‘this proposition is false’ is false, then ‘this proposition is false’ is true.

Here, the subjects of the theses are one and the same, but the implications no longer seem called for, as is made clear if we substitute the symbol P for ‘this proposition is false’. The flavor of paradox has disappeared: it only existed so

long as ‘this proposition is false’ seemed to be implied by or to imply ‘this proposition is true’; as soon as the subject is unified, both the paradoxes break down.

We cannot avoid the issue by formulating the liar paradox as a generality. The proposition ‘I always lie’ can simply be countered by ‘you lie sometimes (as in the case ‘I always lie’), but sometimes you speak truly’; it only gives rise to double paradox in indicative form. Likewise, the proposition ‘all propositions are false’ can be countered by ‘some, some not’, without difficulty.

However, note well, both the said general propositions are indeed self-contradictory; they do produce single paradoxes. It follows that both are false: one cannot claim to ‘always lie’, nor that ‘there are no true propositions’. This is ordinary logical inference, and quite legitimate, since there are logical alternatives.

With regard to those alternatives. The proposition ‘I never lie’ is not in itself inconsistent, except for the person who said ‘I always lie’ intentionally. The proposition ‘all propositions are true’ is likewise not inconsistent in itself, but is inconsistent with the logical knowledge that some propositions are inconsistent, and therefore it is false; so, in this case only the contingent ‘some propositions are true, some false’ can be upheld.

### **3. More on the Liar paradox**

Once we grasp that the meaning of words is their intention, singly and collectively – the solution of the liar paradox becomes very obvious. Self-reference is meaningless, because – an intention cannot intend itself, for it does not yet exist; an intention can only intend something that already exists, e.g. another intention directed at some third thing.

In view of this, the proposition “this proposition is false” is meaningless, and so is the proposition “this proposition is true”. Both may freely be declared equally true and false, or neither true nor false – it makes no difference in their case, because the words “this proposition” refer to nothing at all<sup>11</sup>.

Although the words used in these sentences are separately meaningful, and the grammatical structure of the sentences is legitimate – the words’ collective lack of content implies their collective logical value to be nil. Self-reference is syntactically cogent, but semantically incoherent. It is like circular argument, up in the air, leading nowhere specific.

Regarding the exclusive proposition “*Only* this proposition is true”, it implies both: “*This* proposition is true” and “*All other* propositions are false” – i.e. it is equivalent to the exceptive proposition “*All* propositions *but this* one are false”. The latter is often claimed by some philosopher; e.g. by those who say “all is illusion (except this fact)”.

My point here is that such statements do not only involve the fallacy of self-reference (i.e. “this proposition”). Such statements additionally involve a reference to “all others” which is open to criticism, because:

- To claim knowledge of “all other propositions” is a claim to *omniscience*, a pretense that one knows everything there is to know, or ever will be. And generally, such statements are made without giving a credible justification, though in contradiction to all prior findings of experience and reason.
- Surely, *some* other propositions are in fact regarded and admitted as true by such philosophers. They are generally rather talkative, even verbose – they do not consistently *only* say that one statement and refuse to say anything else.

---

<sup>11</sup> See *Future Logic*, chapter 32.2.



- And of course, formally, if “this” is meaningless (as previously shown), then “all others”, which means “any other *than this*” is also meaningless!

The liar paradox, by the way, is attributed to the ancient Greeks, either Eubulides of Miletus (4<sup>th</sup> Cent. BCE) or the earlier Epimenides of Crete (6<sup>th</sup> Cent. BCE). I do not know if its resolution was evident to these early logicians, but a (European?) 14<sup>th</sup> Cent. CE anonymous text reportedly explained that the Liar’s statement is neither true nor false but simply meaningless. Thus, this explanation is historically much earlier than modern logic (Russell et alia, though these late logicians certainly clarified the matter).<sup>12</sup>

#### 4. The utility of paradoxes

A (*single*) *paradoxical proposition* has the form “if P, then notP” or “if notP, then P”, where P is any form of proposition. It is important to understand that *such propositions are logically quite legitimate within discourse: a (single) paradox is not a contradiction*. On the other hand, a *double paradox*, i.e. a claim that both “if P, then notP” and “if notP, then P” are true in a given case of P, is indeed a contradiction.

The law of non-contradiction states that the conjunction “P and notP” is logically impossible; i.e. contradictory propositions cannot both be true. Likewise, the law of the excluded middle states that “notP and not-notP” is logically unacceptable. The reason for these laws is that such situations of antinomy put us in a cognitive quandary – we are left with no way out of the logical difficulty, no solution to the inherent problem.

On the other hand, single paradox poses no such threat to rational thought. It leaves us with a logical way out – namely, denial of the antecedent (as self-contradictory) and

---

<sup>12</sup>

See *Future Logic*, chapter 63, sections 3 and 6.

affirmation of the consequent (as self-evident). The proposition “if P, then notP” logically implies “notP”, and the proposition “if notP, then P” logically implies “P”. Thus, barring double paradox, *a proposition that implies its own negation is necessarily false, and a proposition that is implied by its own negation is necessarily true.*

It follows, by the way, that the conjunction of these two hypothetical propositions, i.e. double paradox, is a breach of the law of non-contradiction, since it results in the compound conclusion that “P and notP are both true”. Double paradox also breaches the law of the excluded middle, since it equally implies “P and notP are both false”.

These various inferences may be proved and elucidated in a variety of ways:

- Since a hypothetical proposition like “if x, then y” means “x and not y is impossible” – it follows that “if P, then notP” means “P and not notP are impossible” (i.e. P is impossible), and “if notP, then P” means “notP and not P are impossible” (i.e. notP is impossible). Note this explanation well.

We know that the negation of P is the same as notP, and the negation of notP equals P, thanks to the laws of non-contradiction and of the excluded middle. Also, by the law of identity, repeating the name of an object does not double up the object: it remains one and the same; therefore, the conjunction “P and P” is equivalent to “P” and the conjunction “notP and notP” is equivalent to “notP”.

Notice that the meaning of “if P, then notP” is “(P and not notP) is impossible”. Thus, although this implies “notP is true”, it does *not* follow that “if notP is true, P implies notP”. Similarly, *mutatis mutandis*, for “if notP, then P”. We are here concerned with strict implication (logical necessity), not with so-called material implication.

The reason why this strict position is necessary is that in practice, truth and falsehood are contextual – most of what we believe true today might tomorrow turn out to be false, and vice-versa. On the other hand, logical necessity or impossibility refer to a much stronger relation, which in principle once established should not vary with changes in knowledge context: it applies to *all* conceivable contexts.

- Since a hypothetical proposition like “if x, then y” can be recast as “if x, then (x and y)” - it follows that “if P, then notP” equals “if P, then (P and notP)”, and “if notP, then P” equals “if notP, then (notP and P)”. In this perspective, a self-contradictory proposition implies a contradiction; since contradiction is logically impermissible, it follows that such a proposition must be false and its contradictory must be true. This can be expressed by way of apodosis, in which the laws of thought provide the categorical minor premise, making it possible for us to exceptionally draw a categorical conclusion from a hypothetical premise.

If P, then (P and notP)  
 but: not(P and notP)  
 therefore, not P

If notP, then (notP and P)  
 but: not(notP and P)  
 therefore, not notP

- We can also treat these inferences by way of dilemma, combining the given “if P, then notP” with “if notP, then notP” (the latter from the law of identity); or likewise, “if notP, then P” with “if P, then P”. This gives us, constructively:

If P then notP – and if notP then notP  
 but: either P or notP  
 therefore, notP

If notP then P – and if P then P  
 but: either notP or P  
 therefore, P

Paradox sometimes has remote outcomes. For instance, suppose Q implies P, and P implies notP (which as we saw can be rewritten as P implies both P and notP). Combining these propositions in a syllogism we obtain the conclusion “if Q, then P and notP”. The latter is also a paradoxical proposition, whose conclusion is “notQ”, even though the contradiction in the consequent does not directly concern the antecedent. Similarly, non-exclusion of the middle may appear in the form “if Q, then neither P nor notP”. Such propositions are also encountered in practice.

It is interesting that these forms, “Q implies (P and notP), therefore Q is false” and “Q implies (not P and not notP), therefore Q is false”, are the arguments implicit in our application of the corresponding laws of thought.

When we come across an antinomy in knowledge, we dialectically seek to rid ourselves of it **by finding and repairing some *earlier* error(s) of observation or reasoning**. Thus, paradoxical argument is not only a derivative of the laws of thought, but more broadly the very way in which we regularly apply them in practice.

That is, the dialectical process we use following discovery of a contradiction or an excluded middle (or for that matter a breach of the law of identity) means that we believe that:

**Every apparent occurrence of antinomy is, in reality, an illusion.**

It is an illusion *due to paradox*, i.e. it means that *some of the premise(s)* that led to this apparently contradictory or middle-excluding conclusion are in error and in need of correction. The antinomy is never categorical, but hypothetical; it is a sign of and dependent on some wrong previous supposition or assumption. The apparent antinomy serves knowledge by revealing some flaw in its totality, and encouraging us to review our past thinking.

Contradiction and paradox are closely related, but not the same thing. Paradox (i.e. single not double paradox) is not equivalent to antinomy. We may look upon them as cognitive difficulties of different degrees. In this perspective, whereas categorical antinomy would be a dead-end, blocking any further thought—paradox is a milder (more hypothetical) degree of contradiction, one open to resolution.

We see from all the preceding (and from other observations below) the crucial role that paradox plays in logic. The logic of paradoxical propositions does not merely concern some far out special cases like the liar paradox. It is an essential tool in the enterprise of knowledge, helping us to establish the fundamentals of thought and generally keeping our thinking free of logical impurities.

Understanding of the paradoxical forms is not a discovery of modern logic<sup>13</sup>, although relatively recent (dating perhaps from 14<sup>th</sup> Cent. CE Scholastic logic).

---

<sup>13</sup> For instance, Charles Pierce (USA, 1839-1914) noticed that some propositions imply all others. I do not know if he realized this is a property of self-contradictory or logically impossible propositions; and that self-evident or necessary propositions have the opposite property of being implied by all others. I suspect he was thinking in terms of material rather than strict implication.

## 6. The Liar paradox (redux)

*From A Fortiori Logic, appendix 7.4.*

I dealt with the Liar paradox previously, in my *Future Logic*<sup>14</sup>, but now realize that more needs to be said about it. This paradox is especially difficult to deal with because it resorts to several different discursive ‘tricks’ simultaneously.

### 1. First approach

The statement “This proposition is false” looks conceivable offhand, until we realize that if we assume it to be true, then we must admit it to be indeed false, while if we assume it to be indeed false, then we must admit it to be true – all of which seems unconscionable. Obviously, there is a contradiction in such discourse, since nothing can be both true and false. But the question is: just what is causing it and how can it be resolved? We are not ‘deducing’ the fact of contradiction from a ‘law of thought’ – we are ‘observing’ the fact through our rational faculty. We cannot, either, ‘deduce’ the resolution of the contradiction from a ‘law of thought’ – we have to analyze the problem at hand very closely and creatively propose a satisfying solution to it, i.e. one which indeed puts our intellectual anxiety to rest. As we shall see, this is by no means a simple and straightforward matter.

The proposition “This proposition is false” is a double paradox, because: *if it is true, then it is false; and if it is*

---

<sup>14</sup> See there, chapter 32.2. (See also *Ruminations* 5.1.)

*false, then it is true.* Notice the circularity from true to false and from false to true. The implications we draw from the given proposition seem unavoidable at first sight. But we must begin with wonder *how we know these implications* (the two if–then statements) to be true. How do we know that “it is true” implies “it is false,” and that “it is false” implies “it is true”? Apparently, we are not ‘deducing’ these implications from some unstated proposition. We are, rather, using ad hoc rational insight of some sort – i.e. in a sense directly ‘perceiving’ (intellectually cognizing) the implications of the given proposition. But such rational insight, though in principle reliable, is clearly *inductive*, rather than deductive, in epistemological status. That is to say, it is trustworthy until and unless it is found for some reason to be incorrect. This means, there may be one or more errors in our thinking, here; it is not cast in stone. And indeed, there must be some error(s), since it has led to double paradox. Therefore, we must look for it.

Perhaps use of the pronoun “it” is a problem, for it is a rather vague term. Let us therefore ask the question: *more precisely what does the pronoun “it” refer us to, here?*

At first sight, the “it” in “if it is true, then it is false; and if it is false, then it is true” refers to *the whole given statement*, “This proposition is false.” In that event, we must reword the double paradox as follows: *if ‘this proposition is false’ is true, then ‘this proposition is false’ is false; and if ‘this proposition is false’ is false, then ‘this proposition is false’ is true.* Here, the subject of the two if–then statements is more clearly marked out as “this proposition is false,” and so remains constant throughout. But this clarification reveals abnormal changes of predicate, from “true” to “false” and from “false” to “true,” which cannot be readily be explained. Normally, we would say: if ‘this proposition is false’ is true, then ‘this proposition is false’ is *true*; and if ‘this proposition is false’ is false, then ‘this proposition is false’ is *false*. The reason

we here reverse the predicates is that we consider the original proposition, “this proposition is false,” as instructing such reversal.

However, whereas a proposition of the form “‘this proposition is false’ is true” is readily interpretable in the simpler form “this proposition is false,” a proposition of the form “‘this proposition is false’ is false” cannot likewise be simplified. How would we express the double negation involved? As “this proposition is true”? Clearly, the meaning of the latter is not identical to that of the former, since the subject “this proposition” refers to different propositions in each case. So, the formulation of the liar paradox in full form, i.e. as “if ‘this proposition is false’ is true, then ‘this proposition is false’ is false; and if ‘this proposition is false’ is false, then ‘this proposition is false’ is true,” does not make possible the reproduction of the initial formula expressed in terms of the pronoun “it.”

## 2. Second approach

Let us therefore try something else. If the pronoun “it” refers to *the term* “this proposition”, then the double paradox should be reformulated as follows: *if ‘this proposition’ is true, then ‘this proposition’ is false; and if ‘this proposition’ is false, then ‘this proposition’ is true.* But doing that, we see that in each of these two if-then statements, though the subject (“this proposition”) remains constant throughout, the predicate (“true” or “false,” as the case may be) is not the same in the consequent as it was in the antecedent. There is no logical explanation for these inversions of the predicate. Normally, the truth of a proposition P does not imply its falsehood or vice versa.

We might be tempted to use the given “This proposition is false” as a premise to justify the inference from the said antecedents to the said consequents. We might try to formulate two apodoses, as follows:



If this proposition is true, then it is false  
(hypothesis),  
and this proposition is false (given);  
therefore, this proposition is true (putative  
conclusion).

If this proposition is false, then it is true  
(hypothesis),  
and this proposition is false (given);  
therefore, this proposition is true (putative  
conclusion).

Obviously, in the first case we have invalid inference, in that we try to deny the antecedent to deny the consequent, or to affirm the consequent to affirm the antecedent. In the second case, the putative conclusion does follow from the premises; but we can still wonder where the major premise (the hypothetical proposition) came from, so we are none the wiser. So, this approach too is useless – i.e. it proves nothing.

Alternatively, we might try formulating the following two syllogisms:

This proposition is false (given),  
and this proposition is true (supposition);  
therefore, this proposition is false (putative  
conclusion).

This proposition is false (given).  
 and this proposition is false (supposition);  
 therefore, this proposition is true (putative  
 conclusion).

Clearly, these arguments are not quite syllogistic in form; but they can be reworded a bit to produce syllogisms. The first two premises would then yield the conclusion “there is a proposition that is true and false” (3/RRI), which is self-contradictory (whence, one of the premises must be false); the second two premises, however, being one and the same proposition, would yield no syllogistic conclusion other than “there is a proposition that is false and false” (3/RRI), which is self-evident (and trivial). But these are not the conclusions we seek, which must concern “this proposition” and not merely “some proposition.”

A better approach is to look upon the latter two arguments as follows. In the first case, the premises “this proposition is false” (given) and “this proposition is true” (supposition) seem to together imply “this proposition is both true and false;” and the latter paradoxical conclusion in turn indeed suggests that “this proposition is false,” since contradiction is impossible. And in the second case, the premises “this proposition is false” (given) and “this proposition is false” (supposition) agree with each other that “this proposition is false,” and so this is their logical conclusion. Since both arguments conclude with “this proposition is false,” the latter must be the overall conclusion.

However, the latter result is not as conclusive as it seems, because upon closer scrutiny it is obvious that “this proposition is false” and “this proposition is true” do not refer to the same subject, since the predicate changes. The first “this proposition” refers to the proposition “this proposition is false” and the second “this proposition”

refers to the proposition “this proposition is true.” So, these two propositions in fact have different subjects as well as different predicates (viz. false and true, respectively). The subjects superficially look the same, because they are verbally expressed in identical words; but their underlying intent is not the same, since they refer to significantly different propositions (propositions with manifestly different, indeed contradictory, predicates). This means that when the predicate changes, the subject effectively changes too. When the predicate is “true,” the subject means one thing; and when the predicate is “false,” the subject means something else. Although the words “this proposition” are constant, their underlying intent varies. That is to say, the term “this proposition” does not have a uniform meaning throughout, and therefore cannot be used as a basis for the inferences above proposed.

### 3. Third approach

Let us now try another angle. If we examine our initial reasoning in terms of the pronoun “it” more carefully, we can see what is really happening in it. Given that ‘this proposition is false’ is true, we can more briefly say: ‘this proposition is false.’ Also, given ‘this proposition is false’ is false, we can by negation educe that ‘this proposition is *not* false’ is true, which means that ‘this proposition is true’ is true, or more briefly put: ‘this proposition is true’<sup>15</sup>. In this way, we *seem* to argue, regarding the subject “this proposition is false,” from ‘it is true’ to ‘it is false’, and from ‘it is false’ to ‘it is true’. But in fact, the use of the pronoun “it” or the term “this proposition” as abbreviated

---

<sup>15</sup> Some logicians have tried to deal with the liar paradox by denying that true and false are contradictory terms, i.e. that not-true = false and not-false = true. Such a claim is utter nonsense; the attempt to shunt aside the laws of non-contradiction of the excluded middle so as to resolve a paradox is self-contradiction in action.

subject is a sleight of hand, for the underlying subject changes in the course of the second transition (that ending in “this proposition is true”). When abbreviation is used throughout, we seem to be talking about one and the same proposition throughout as being both true and false. But seeing that this is based on hidden equivocation, the paradoxes disappear.

It is interesting to note that when the reasoning is viewed more explicitly like that, the proposition “this proposition is true” also becomes paradoxical! We can argue: if ‘this proposition is true’ is true, then obviously ‘this proposition is true’. And: if ‘this proposition is true’ is false, then its contradictory ‘this proposition is *not* true’ must be true, which means that ‘this proposition is false’ is true, i.e. more succinctly: ‘this proposition is false’. Here, superficially, there seems to be no paradox, because we seem to argue, regarding the subject “this proposition is true,” from ‘it is true’ to ‘it is true’, and from ‘it is false’ to ‘it is false’. But if we look at the final conclusion, viz. “this proposition is false,” we see that it *corresponds to* the liar paradox!<sup>16</sup> And here again, the explanation of the double paradox is that the apparent subject “it” or “this proposition” changes significance in the course of drawing the implications.

Notice that, in both these lines of reasoning, the first leg is ordinary self-implication, mere tautology, while the second leg is the operative self-contradiction, the paradox. If the given proposition (whether “this proposition is false” or “this proposition is true”) is true, we merely repeat the proposition as is (without need to add the predication “is true”). But if the given proposition is false, we cannot drop

---

<sup>16</sup> That ‘this proposition is true’ is implicitly (if only potentially) as paradoxical as ‘this proposition is false’ is, so far as I know, a new discovery. Note well how *both* paradoxes occur through quite ordinary eductions: viz. if ‘P is Q’ is affirmed, then P is Q; and if ‘P is Q’ is denied, then ‘P is not Q’ is affirmed, then P is not Q (where P stands for ‘this proposition’, and Q for ‘false’ or ‘true’ as the case may be).

the additional predication (i.e. “is false”) without changing the original proposition. Thus, we could say that the two propositions, “this proposition is false” or “this proposition is true,” present no problem when taken as true; and it is only when they are hypothetically taken as false that the problem is created. So, we could say that the way out of the liar paradox (and its positive analogue) is simply to accept the two claims as true, and not imagine them to be false!

We could furthermore, if we really want to, argue that “this proposition is false” and “this proposition is true” differ in that the former explicitly appears to put itself in doubt whereas the latter does not do so. On this basis, we could immediately reject the former and somewhat accept the latter, even while admitting that the latter is equally devoid of any useful information. That is to say, since the former appears ‘more paradoxical’ than the latter, the latter is to be preferred *in extremis*. But this, note well, ignores the equally insurmountable difficulties in it. It is better to resolutely reject both forms as vicious constructs.

#### 4. Fourth approach

To grasp the illusoriness of the liar paradox, it is important to realize that the two forms, “this proposition is false” and “this proposition is true,” are *not* each other’s contradictory; and that, in fact, neither of them *has* a contradictory! This is a *logical anomaly*, a fatal flaw in the discourse of the liar paradox; for in principle, every well-formed and meaningful proposition is logically required to have a contradictory. If a propositional form lacks a contradictory form, it cannot be judged true or false, for such judgment depends on there being a choice. We do not even have to limit our propositions to the predicates “true” or “false” – any predicate X and its negation not-X would display the same property given the same said subject. That is, “this proposition is X” and “this proposition is not-X”

are *not* each other's contradictory, and are therefore *both* equally deprived of contradictory.

We could, of course, remark that "this proposition is X" can be denied by "*that* proposition (i.e. the preceding one) is not X," or even introduce *a symbol* for the original proposition in the new proposition. In such case, although the subjects would be verbally different, their intents would surely be the same. But the form "that proposition is not X" is more akin to the form "'this proposition is X' is not X," in which the whole original proposition is given the role of subject and its predicate is given the role of predicate. However, though these two forms are somewhat equivalent in meaning to each other and to the original proposition, their logical behavior patterns are not identical with that of the original proposition, as we have already seen. The fact remains that "this proposition is not X" is not the contradictory of "this proposition is X."

Clearly, any proposition involving the special subject "this proposition" exhibits a very unusual property, and may be dismissed on that basis alone. The reason why such a proposition lacks a contradictory is that its subject refers to the proposition *it happens to be in*, and that proposition is evidently different when the predicate in it is the term "false" and when it is the term "true" (or more generally, any pair of predicates 'X' and 'not-X'). When the predicate changes, *so does the subject*; so the subject cannot be pinned-down, it is variable, it is not constant as it should be. The term "this proposition" has a different reference in each case, which depends on the predicate; consequently, each subject can only be associated with one predicate and never with the other (i.e. its negation).

From this we see that when at the beginning we thought, looking upon the statement "This proposition is false," that if we take it at its word, then it is must be regarded as false, and so we have to prefer to it "This proposition is not

false,” i.e. “This proposition is true,” and so forth, we did not realize that we were in fact, due to the ambiguity inherent in the term “This proposition” or “it,” changing its meaning at every turn. This change of meaning passes by unnoticed, because the term used is by its very nature not fixed. The pronouns “this” and “it” can be applied to anything and its opposite without such change of meaning being verbally signaled in them. They are not permanently attached to any object, but are merely contextual designations. In the technical terminology of linguistics, they are characterized as ‘deictic’ or ‘indexical’.

Thus, it appears that the liar paradox arises, however we understand its terms, as a result of some sort of equivocation in the subject. Although we seem superficially to refer to one and the same subject in the antecedent and consequent of our if–then reasoning, there is in fact a covert change of meaning which once we become aware of it belies the initial appearance of contradiction. The suggested impossible implications are thus put in doubt, made incredible. The contradictions apparently produced are thus defused or dissolved, by virtue of our inability to make them stick.

## 5. Fifth approach

Another, and complementary, way to deal with the liar paradox is to point out *the logical difficulty of self-reference*. This is a tack many logicians have adopted, including me in my first foray into this topic in *Future Logic*. The argument proposed here is that the term “this proposition” refers to an object (viz. “This proposition is false” or “This proposition is true”) which includes the term itself. A finger cannot point at itself, and “this” is the conceptual equivalent of a finger. Effectively, the expression “this” has no content when it is directed at itself or at a sentence including it. It is empty, without substance.

It is as if nothing is said when we indulge in such self-reference.

Thus, “This proposition is X” (where X stands for false, or true, or indeed anything) is in fact meaningless; and a meaningless sentence cannot be true or false. Such a sentence can reasonably be described as neither true nor false, without breach of the law of the excluded middle, because neither of these logical evaluations is applicable to meaningless sentences. “This proposition is false” looks meaningful because its four constituents (i.e. “this,” “proposition,” “is” and “false”) are separately normally meaningful. But in this particular combination, where one of the elements (viz. “this”) does not refer to anything already existent, the sentence is found to be meaningless.

The apparent contradictions that self-reference produces help us to realize its meaninglessness. And it is through the intellectual realization of the meaninglessness of self-reference that we explain away and annul the apparent contradictions. On this basis, we can say that even though the sentence “This proposition is true” does not at first sight give rise to any paradox (as people think: “if it is true, it is true; and if it is false, it is false”), nevertheless, since it involves self-reference as much as “This proposition is false,” it is equally meaningless and cannot be characterized as true or false. In fact, as I have shown above, “This proposition is true” does also give rise to double paradox.

Someone might object: What about the propositions: “this statement is self-referential” and “this statement is not self-referential”? Surely, we can say that these are meaningful and that the former is true while the latter is false! The retort to that objection is that the two propositions “this proposition refers to itself” and “this proposition does not refer to itself” are *not* mutual contradictories, because (just like in the liar paradox) their subjects differ radically, each



referring to the proposition it is in and not to the other. Thus, while the positive version may seem more self-consistent than the negative one, and therefore to be preferred *in extremis*, they are in fact both fundamentally flawed, because (just like in the liar paradox) neither of them *has* a contradictory, and without the logical possibility of negating a discourse it is impossible to judge whether it is right or wrong.<sup>17</sup>

## 6. Sixth approach

Not long after the preceding reflections, I happened to come across another interesting example of paradoxical self-reference, namely “Disobey me!”<sup>18</sup> This involves the ‘double bind’ – *if I obey it, I disobey it and if I disobey it, I obey it*. To resolve this paradox, we need to first put the

---

<sup>17</sup> Another objection (which was actually put to me by a reader) would be propositions like “this statement has five words” and “this statement has six words” – even though they contain the demonstrative “this,” the former looks true and the latter false! Here, we might in reply point out that though the propositions “this statement has five words” and “this statement does not have five words,” seem to mean opposite things, they cannot be contradictories, since both appear true. Also compare: “this statement has five words” and “this statement does have five words” – the former is true while the latter is false, though both *mean* essentially the same. Clearly, the behavior of these propositions is far from normal, due to their unusual dependence on the wording used in them. On one level, we get the message of the proposition and count the number of words in it, and then check whether this number corresponds to the given number: if yes, the proposition is judged ‘true’, and if no, it is judged ‘false’. But at the same time, we have to be keep track of the changing reference of the demonstrative “this,” which complicates matters as already explained, and additionally in this particular context we must beware of the impact of wording. The Kneales give “What I am now saying is a sentence in English” as an example of “harmless self-reference” (p. 228).

<sup>18</sup> I found this example in Robert Maggiori’s *La philosophie au jour le jour*; the author does not say whether it is his own invention or someone else’s (p. 438).

statement in more precise form, say: “you must disobey this command!” We can then disentangle the knot by realizing that the order being given *has outwardly imperative form but inwardly lacks content*. It does not define a specific, concrete action that is to be done or not-done. If we wished to obey it, or to disobey it, we would not know just what we are supposed to do or not-do! It is therefore an order that can neither be obeyed nor be disobeyed. Ruminating on this case led me to what I now believe is the trump card, which convincingly finalizes the resolution of the liar paradox, even as the preceding reflections all continue to be relevant.

It occurred to me then that this is precisely the problem with the liar paradox. It says “this proposition is false” – but *it does not tell us anything about the world that can be judged as true or false*. A ‘proposition’ is a statement that makes some claim about the world. If the statement makes no such claim, if it ‘proposes’ nothing, it cannot be logically assessed as true or false. If it refers to nothing – whether physical, mental or spiritual, perceptual, intuitive or conceptual – it has no meaning. A meaningless statement does not qualify as a ‘proposition’. The attributes of ‘true’ or ‘false’ are not ordinary predicates, like ‘white’ or ‘black’, which can be attached to any subject and then judged to be truly or falsely attached. The attributes of ‘true’ or ‘false’ require a precise claim to be made before they can at all be used.

The truth of this explication can be seen with reference to the ‘propositional forms’ used in logic theory. Take, for example, “All X are Y.” Such a propositional form cannot be judged true or false because it manifestly has no content. Only when such an abstraction is given some specific content, such as “All men are mortal,” can we begin to ask whether it is true or false. A propositional form is too vague to count as a proposition. It does not tell us anything about the world, other than implying that there are (or even just

that there may be) concrete propositions which have this form. Just as we cannot disobey or even obey an imperative without content, so we cannot judge a purely formal expression true or false.

The same applies to the liar paradox: like a formal proposition, it has no concrete content, and therefore cannot be judged true or false. The liar paradox has no content partly due to its having a self-referential subject (“this proposition”). But the truth is, *even if its subject was not self-referential*, it would still have insufficient content. This is so, because its predicate “false” (and likewise its opposite, “true”) is not an ordinary predicate; it is more like a formal predicate. It can only be used if another, more concrete predicate has already been proposed for the subject at hand. For example, “this proposition is interesting” could be judged true or false (if it was not self-referential) because it already has a predicate (viz. “interesting”). Thus, the problem with the liar paradox is not only the self-reference it involves but also its lack of a predicate more concrete than the logical predicate “false” (or “true”).

All this illustrates how the ‘laws of thought’ are not axioms in the sense of top premises in the knowledge enterprise from which we mechanically derive other premises. Rather the expression ‘laws of thought’ refers to recurring insights which provide us with some intellectual guidance but cannot by themselves determine the outcome. The individual in pursuit of knowledge, and in particular the logician, is *driven* by the obviousness or by the absurdity of a situation to look for creative solutions to problems. He or she must still think of possible solutions and test them.

## 7. The Russell paradox (early)

*From Future Logic 32, 42 & 45, and Ruminations 5.*

### 1. Self-membership

With regard to the issue of self-membership, more needs to be said. Intuitively, to me at least, the suggestion that something can be both container and contained is hard to swallow.

Now, self-membership signifies that a nominal is a member of an exactly identical nominal. Thus, that all X are X, and therefore members of “X”, does not constitute self-membership; this is merely the definition of membership in a first order class by a non-class.

We saw that, empirically, at least with ordinary examples, “X” (or the class of X) is never itself an X, nor therefore a member of “X”. For example, “dogs” is not a dog, nor therefore a member of “dogs”.

I suggested that this could be generalized into an inductive postulate, if no examples to the contrary were forthcoming. My purpose here is to show that all apparent cases of self-membership are illusory, due only to imprecision of language.

That “X” is an X-class, and so a member of “X-classes”, is not self-membership in a literal sense, but is merely the definition of membership in a second order class by a first order class. For example, “dogs” is a class of dogs, or a member of “classes of dogs”, or member of the class of classes of dogs.

Nor does the formal inference, from all X are X, that all X-classes are X-classes, and so members of “X-classes” (or the class of classes of X), give us an instance of what we strictly mean by self-membership; it is just tautology. For example, all dog-classes are members of “classes of dogs”. Claiming that an X-class may be X, and therefore a member of “X”, is simply a wider statement than claiming that “X” may be X, and not only seems equally silly and without empirical ground, but would in any case not formally constitute self-membership. For example, claiming “retrievers” is a dog.

As for saying of any X that it *is* “X”, rather than a member of “X”; or saying that it *is* some other X-class, and therefore a member of “X-classes” — such statements simply do not seem to be in accord with the intents of the definitions of classes and classes of classes, and in any case are not self-membership.

The question then arises, is “X-classes” itself a member of “X-classes”? The answer is, no, even here there is no self-membership. The impression that “X-classes” might be a member of itself is due to the fact that it concerns X, albeit less directly so than “X” does. For example, dog-classes refers to “retrievers”, “terriers”, and even “dogs”; and thus, though only indirectly, concerns dogs.

However, more formally, “X-classes” does not satisfy the defining condition for being a member of “X-classes”, which would be that ‘all X-classes are X’ — just as: “X” is a member of “X-classes”, is founded on ‘all X are X’. As will now be shown, this means that the above impression cannot be upheld as a formal generality, but only at best as a contingent truth in some cases; as a result, all its force and credibility disappear.

If we say that *for any and every* X, all X-classes are X, we imply that for all X, “X” (which is one X-class) is X; but we have already adduced empirical cases to the contrary;

so, the connection cannot be general and formal. Thus, we can only claim that perhaps *for some X*, all X-classes are X; but with regard to that eventuality, no examples have been adduced.

Since we have no solid grounds (specific examples) for assuming that “X” or “X-classes” is ever a member of itself, and the suggestion is fraught with difficulty; and we only found credible examples where they were not members of themselves — we are justified in presuming, by generalization, that: *no class of anything, or class of classes of anything, is ever a member of itself.*

I can only think of one possible exception to this postulate, namely: “things” (or “things-classes”). But I suspect that, in this case, rather than saying that the class is a member of itself, we should regard the definition of membership as failing. That is, though this summum genus *is* a thing, it is not ‘a member of’ anything.

## 2. The Russell paradox

The Russell Paradox is modern example of double paradox, discovered by British logician Bertrand Russell.

He asked whether the class of “all classes which are not members of themselves” is or not a member of itself. If “classes not members of themselves” is not a member of “classes not members of themselves”, then it is indeed a member of “classes not members of themselves”; and if “classes not members of themselves” is a member of “classes not members of themselves”, then it is also a member of “classes which are members of themselves”. Thus, we face a contradiction either way.

In contrast, the class of “all classes which are members of themselves” does not yield a similar difficulty. If “self-member classes” is not a member of “self-member classes”, then it is a member of “classes not members of

themselves”; but if “self-member classes” is a member of “self-member classes”, no antinomy follows. Hence, here we have a single paradox coupled with a consistent position, and a definite conclusion can be drawn: “self-member classes” is a member of itself.

Now, every absurdity which arises in knowledge should be regarded as an opportunity for advancement, a spur to research and discovery of some previously unknown detail. So, what is the hidden lesson of this puzzle?

As I will show, the Russell Paradox proceeds essentially from an equivocation; it is more akin to the sophism of the Barber paradox, than to that of the Liar paradox. For *whether self-membership is possible or not, is not the issue*. Russell believed that some classes, like “classes” include themselves; though I disagree with that, my disagreement is not my basis for dissolving the Russell paradox. For it is not the concept of self-membership which results in a two-way inconsistency. It is the concept of non-self-membership which does so; and everyone agrees that at least some (if not all, as I believe) classes do not include themselves: for instance, “dogs” is not a dog.

What has stumped so many logicians with regard to the Russell paradox, was the assumption that we can form concepts at will, if we but formulate a verbal definition. But this viewpoint is without justification. The words must have a demonstrable meaning; in most cases, they do; but in some cases, they are isolated or pieced together without attention to their intrinsic structural requirements. We cannot, for instance, use the word ‘greater’ without specifying ‘than what?’; many words are attached, and cannot be reshuffled at random. The fact that we commonly, in everyday discourse, use words loosely, to avoid boring constructions, does not give logicians the same license.

### 3. Impermutability

The solution to the problem is so easy, it is funny, though I must admit I was quite perplexed for a while. It is simply that: *propositions of the form ‘X (or “X”) is (or is not) a member of “Y” (or “Y-classes”)’ cannot be permuted. The process of permutation is applicable to some forms, but not to all forms.*

a. In some cases, where we are dealing with relatively simple relations, the relation can be attached to the original predicate, to make up a new predicate, in an ‘S is P’ form of proposition, in which ‘is’ has a strictly classificatory meaning. Thus, ‘X is-not Y’ is permutable to ‘X is nonY’, or ‘X is something which is not Y’; ‘X has (or lacks) Y-ness’ is permutable to ‘X is a Y-ness having (or lacking) thing’; ‘X does (or does not do) Y’ is permutable to ‘X is a Y-doing (or Y-not-doing) thing’. In such cases, no error arises from this artifice.

But in other cases, permutation is not feasible, because it falsifies the logical properties of the relation involved. We saw clear and indubitable examples of this in the study of modalities.

For instance, the form ‘X can be Y’ is not permutable to ‘X is something capable of being Y’, for the reason that we thereby change the subject of the relation ‘can be’ from ‘X’ to ‘something’, and also we change a potential ‘can be’ into an actual ‘is (capable of being)’. As a result of such verbal shenanigans, formal errors arise. Thus, ‘X is Y, and all Y are capable of being Z’ is thought to conclude ‘X is capable of being Z’, whereas in fact the premises are quite compatible with the contradictory ‘X cannot be Z’, since ‘X can become Z’ is a valid alternative conclusion, as we saw earlier.

It can likewise be demonstrated that ‘X can become Y’ is not permutable to ‘X is something which can become Y’, because then the syllogism ‘X is Y, all Y are things which



can become Z, therefore X is something which can become Z' would seem valid, whereas its correct conclusion is 'X can be or become Z', as earlier seen. Thus, modality is one kind of relational factor which is not permutable. Even though we commonly say 'X is capable or incapable of Y', that 'is' does not have the same logical properties as the 'is' in a normal 'S is P' proposition.

b. The Russell Paradox reveals to us the valuable information that the copula 'is a member (or not a member) of' is likewise not open to permutation to 'is something which is a member (or not a member) of'.

The original 'is' is an integral part of the relation, and does not have the same meaning as a solitary 'is'. The relation 'is or is not a member of' is an indivisible whole; you cannot just cut it off where you please. The fact that it consists of a string of words, instead of a single word, is an accident of language; just because you can separate its verbal constituents does not mean that the objective relation itself can similarly be split up.

Permutation is a process we use, when possible, to bypass the difficulties inherent in a special relation; in this case, however, we cannot get around the peculiar demands of the membership relations by this artifice. The Russell paradox locks us into the inferential processes previously outlined; it tells us that there are no other legitimate ones, it forbids conceptual short-cuts.

The impermutability of 'is (or is not) a member of' signifies that you cannot form a class of 'self-member classes' or a class of 'non-self-member classes'. These are not terms, they are relations. Thus, the Russell paradox is fully dissolved by denying the conceptual legitimacy of its terms. There is no way for us to form such concepts; they involve an illicit permutation. The connections between the terms are therefore purely verbal and illusory.

The definition of membership is ‘if something *is* X, then it is a member of “X” or ‘if all X *are* Y, then “X” is a member of “Y-classes”’. The Russell paradox makes us aware that the ‘is’ in the condition has to be a normal, solitary ‘is’, it cannot be an ‘is’ isolated from a string of words like ‘is (or is not) a member of’. If this antecedent condition is not met, the consequent rule cannot be applied. In our case, the condition *is not met*, and so the rule does not apply.

c. Here, then, is how the Russell paradox formally arises, step by step. We will signal permutations by brackets like this: {}.

Let “X” signify any class, of any order:

(i) If “X” is a member of “X”, then “X” is {a member of itself}. Call the enclosed portion Y; then “X” is Y, defines self-membership.

(ii) If “X” is not a member of “X”, then “X” is {not a member of itself}. Call the enclosed portion nonY; then “X” is nonY, defines non-self-membership.

Next, apply the general definitions of membership and non-membership to the concepts of Y and nonY we just formed:

(iii) whatever is not Y, is nonY, and so is a member of “nonY”.

(iv) whatever is Y, is not a member of “nonY”, since only things which are nonY, are members of “nonY”.

Now, the double paradox:

(v) if “nonY” is not a member of “nonY”:

— then, by putting “nonY” in place of “X” in (ii), “nonY” is {not a member of itself}, which means it is nonY;

— then, by (iii), “nonY” is a member of “nonY”, which contradicts the starting premise.

(vi) if “nonY” is a member of “nonY”:

— then, by putting “nonY” in place of “X” in (i), “nonY” is {a member of itself}, which means it is Y;

— then, by (iv), “nonY” is not a member of “nonY”, which contradicts the starting premise.

Of all the processes used in developing these arguments, only one is of uncertain (unestablished) validity: namely, permutation of ‘is a member of itself’ to ‘is {a member of itself}’, or of ‘is not a member of itself’ to ‘is {not a member of itself}’. Since all the other processes are valid, the source of antinomy has to be such permutation. Q.E.D.

d. The existence of impermutable relations suggests that we cannot regard all relations as somehow residing *within* the things related, as an indwelling component of their identities. We are pushed to regard some relations, like modality or membership, as bonds standing outside the terms, which are not actual parts of their being.

Thus, for example, that ‘this S can be P’ does not have an ontological implication that there is some actual ‘mark’ programmed in the actual identity of this S, which records that it ‘can be P’. For this reason, the verbal clause {can be P} cannot be presumed to be a unit; there is nothing corresponding to it in the actuality of this S, the potential relation does not cast an actual shadow.

Thus, there must be a reality to ‘potential existence’, outside of ‘actual existence’. When we say that ‘this S can be P’, we consider this potentiality to be P as somehow part of the ‘nature’ of this S. But the S we mean, itself stretches in time, past, ‘present’, and future; it also has ‘potential’ existence, and is wider than the actual S.

The same can be argued for can not, or must or cannot. Thus, natural (and likewise temporal) modalities refer to different degrees, or levels, of existence.

Similarly, the impermutability of membership relations, signifies that they stand external to their terms, leaving no mark on them, even when actual.

It seems like a reasonable position, because if every relation of something to everything else, implied some corresponding trait inside that thing, then each thing in the world would have to contain an infinite number of messages, one message for its relations to each other thing. Much simpler, is to regard relations (at least, those which are impermutable) as having a separate existence from their terms, as other contents of the universe.

#### 4. The Barber paradox

The Barber paradox<sup>19</sup> may be stated as: ‘If a barber shaves everyone in his town who does not shave himself, does he or does he not shave himself? If he does, he does not; if he does not, he does’.

This double paradox arises through confusion of the expressions ‘does not shave himself’ and ‘is shaved by someone other than himself’.

We can divide the people in any town into three broad groups: (a) people who do not shave themselves, but are shaved by others; (b) people who do not shave themselves, and are not shaved by others; (c) people who shave themselves, and are not shaved by others. The given premise is that our barber shaves all the people who fall in group (a). It is tacitly suggested, but not formally implied, that no one is in group (b), so that no one grows a beard or is not in need of shaving. But, in any case, the premise in fact tells us nothing about group (c).

---

<sup>19</sup> This paradox was offered by B. Russell (in his 1918-19 work, *The Philosophy of Logical Atomism*) as an illustration of the Russell paradox; but he did not claim it as his own, saying that it was “suggested” to him by someone else. However, Russell considered that “In this form the contradiction is not very difficult to solve,” because one can simply deny the subject (i.e. say that such a barber does not exist).

Next, let us subdivide each of the preceding groups into two subgroups: (i) people who shave others, and (ii) people who do not shave others. It is clear that each of the six resulting combinations is logically acceptable, since who shaves me has no bearing on whom I can shave. Obviously, only group (i) concerns barbers, and our premise may be taken to mean that our barber is the only barber in town.

Now, we can deal with the question posed. Our barber cannot fall in group (a)(i), because he is not shaved by others. He might fall in group (b)(i), if he were allowed to grow a beard or he was hairless; but let us suppose not, for the sake of argument. This still allows him to fall in group (c)(i), meaning that he shaves himself (rather than being shaved by others), though he shaves others too.

Thus, there is no double paradox. The double paradox only arose because we wrongly assumed that ‘he shaves all those who *do not* shave themselves’ excludes ‘he shaves some (such as himself) who *do* shave themselves’. But ‘*X shaves Y*’ does not formally contradict ‘*X shaves nonY*’; there is no basis for assuming that the copula ‘to shave’ is obvertible, so that ‘*X shaves Y*’ implies ‘*X does not shave nonY*’.

If the premise was restated as ‘he shaves all those *and only* those who do not shave themselves’ (so as to exclude ‘he shaves himself’), we would still have an out by saying ‘he does not shave at all’. If the premise was further expanded and restricted by insisting that ‘he somehow shaves or is shaved’, it would simply be self-contradictory (in the way of a single paradox).

Further embellishments could be made to the above, such as considering people who shave in other towns, or making distinctions between always, sometimes/sometimes-not, and never. But I think the point is made. The lesson learned from the barber ‘paradox’ is that without clear categorizations, equivocations can emerge (such as that

between ‘shaves’ and ‘is shaved’), which give the illusion of double paradox.

## 5. The Master Catalogue paradox

A class may be viewed as an imaginary envelope, which flexibly wraps around all the class’ purported members, however dispersed in place and time, to the exclusion of all other things. The question arises, can the figurative envelope of the class “classes” wrap itself too, or not?

Reviewing the **Russell paradox**<sup>20</sup>, we must conclude that *not all ‘word-objects’ are ‘things’* – measures of things are not themselves to be counted as things. Since classification is an expression of our measurement of things, it cannot itself be counted as a thing. To do so gives rise to a paradox, we should avoid it.

In other words, the problem involved is that the iterative form (“class of classes”) is not identical with the simple form (“class”), except very superficially and verbally – so the former cannot logically be subsumed under the latter. There is a sufficiently significant modification of the subject-predicate relation involved, caused by the iteration of the same term, to exclude the reflex of subsumption. The paradox arising if we do not restrain this impulse is precisely what teaches us to exercise such restraint.

The word ‘things’, note, has many meanings. Sometimes, we intend by it all possible objects of thought. Sometimes, we mean to exclude words from it<sup>21</sup>. Sometimes, we mean to exclude classes; or more narrowly, as just pointed out, classes of classes; ditto, with regard to concepts or to concepts of concepts. Sometimes, the word ‘things’ includes only material objects, whatever their category.

---

<sup>20</sup> See *Future Logic*, chapters 43-45, on class logic.

<sup>21</sup> Though of course, this distinction may be paradoxical, since the word ‘word’ refers to words.

Sometimes, we mean by it ‘entities<sup>22</sup>’ (material, mental or spiritual bodies, or delimited substances, individual cases of which are generally subjects of propositions) in contrast to their ‘properties’ (the predicates of place, time, quality, action, quantity, relation, and so forth). Sometimes, in everyday discourse, we refer to ‘things’ in contrast to ‘persons’ – i.e. ‘things’ here means inanimate or non-volitional entities. And there are yet more senses of the word.

Thus, whenever logicians refer to ‘things’, they ought to try and first make clear just what is to be included under that heading.

Incidentally, even worse than ‘self-membership’ as a concept to swallow, is the notion of “*classes that seem contradictory to what they include*” – the latter seems inconceivable at the outset, at least in verbal appearance! Thus, for instances: “no relationship” is a relationship of sorts; “non-classes” is in a sense a class. There has to be some fallacy involved in such terms, which needs to be clarified. Perhaps the problem is a hyperbole or misnomer?

The answer to this question would be that we are here again dealing with *classes of classes*, and these need not be outwardly consistent with their member classes. Thus, the class of non-relationships still involves a relationship. The class of non-classes is nonetheless a class. The class of empty or null classes does have members. The class of meaningless or self-contradictory classes is itself neither meaningless nor self-contradictory. And so forth.

Bertrand Russell illustrates his paradox with reference to:

- (a) a catalogue of all books that mention themselves, and
- (b) a catalogue of all books that do *not* mention themselves.

---

<sup>22</sup> The word ‘entity’, of course, is sometimes meant more generally, with reference to any existent.

Case (a) presents no problem: the catalogue can list itself without contradicting its own definition; whereas, if it does not list itself, it betrays that definition<sup>23</sup>. Case (b), on the other hand, is a problem: if it does not list itself, in accord with its own definition, it thereby becomes eligible for inclusion in itself<sup>24</sup>; but, if it does indeed list itself, it contradicts its own definition. The latter is the double paradox under discussion.

Now, my first objection would be as follows. The catalogue's title (and even, perhaps, a brief description of its contents, an abstract) could perhaps be listed within the book itself— but such a book would not and cannot include a reproduction of the whole book inside itself (not to mention all the other books it lists or reproduces), for the simple reason that the task would be infinite (a book within a book within a book... etc., or the same in the plural).

The book is therefore not *itself* a member of itself; strictly speaking, only words *about* the book are mentionable in it. The terms inclusion or membership, as used here, then, have a very limited meaning. Thus, the plausibility of Russell's example is very superficial, spurious; he is being fallacious, sophistical, suggesting something impossible.

Moreover, every book "includes itself" in the sense that it consists of whatever contents it has and no more. But if a book is conceived as including a number of other books, defined by some statement (e.g. all English books), the book cannot include itself in the sense that this content is *only part of* itself. This would not only signify infinite regression (a book with other books plus itself in it, the latter in turn with other books plus itself in it, and so forth), and infinite size, but it would constitute a contradiction

---

<sup>23</sup> That is, the catalogue is not eligible for inclusion in itself – but that does not affect its exhaustiveness.

<sup>24</sup> So that, if it is not forthwith included in itself, it can no longer be claimed complete.



within the definition. The book cannot both be all its content and only part of its content.

In this perspective, defining the book as ‘the catalogue of all books that do not include themselves’, the Russell paradox is akin to the liar paradox, since the projected book is an entity that has no finite dimension; it can never be pinned down.

A second objection would be the following. Even if we take Russell’s construct as a mere list of books, defined as ‘the catalogue of all books that do not *mention* themselves’, the definition is absurd, since *it cannot logically be realized*. We simply cannot write a book listing all books that do not mention themselves (Conrad’s *Lord Jim*, Hugo’s *Notre Dame*, etc.), in view of the stated dilemma, that whether we list or not list the book itself in it we are in a contradiction. Therefore, this concept is of necessity a null-class and meaningless.

Logic has not been stumped by the paradox, but has precisely just been taught that the proposed concept is unsound and unusable; it must therefore simply be dropped or at least changed somewhat. There is nothing dramatic in the paradox; it represents one of the functions of Logic. We might try to propose a modified concept, as follows. Perhaps we should instead refer to a library.

(a) Consider a catalogue of all books in a certain library, which is to be placed in that same library. If the book lists itself, it presents no problem. If the book does not mention itself as being in the library, it is simply incomplete and should be expanded; or its title is incorrect and should be modified (“all books but this one”); or it should be left out of the library.

(b) Now, with regard to a catalogue of all books *not* in our library: such a book cannot both mention itself and be put in the library. If we want to keep it in our library, we must erase its mention of itself. If we want it to mention itself,

we must leave it out of the library. These are practical alternatives, which present no problem.

In this perspective, as we seek a practical expression for it, the Russell paradox becomes more akin to the Barber paradox.

## 6. Grelling's paradox

To develop his paradox<sup>25</sup>, Kurt Grelling<sup>26</sup> labels a word 'homological', if it has the quality it refers to (e.g. the word "short" is short, or the word "polysyllabic" is polysyllabic), or 'heterological', if it lacks the quality it refers to (e.g. "long" lacks length, or again "monosyllabic" is not monosyllabic). He then asks whether these two words, themselves, are to be categorized this way or that, arguing:

- If "heterological" is homological, then it is heterological (contradictory predicates).
- If "heterological" is heterological, then it is homological (contradictory predicates).

But it is a misapprehension of the meanings of these words to even try to apply them to themselves. In their case, the references are *too abstract* to have visible or audible concomitants. *Neither* term is applicable to either of them.

Note first that the apparent contradictions in predication either way apply to the word "heterological" only. For, using similar reasoning with regard to the word "homological", although it might seem more consistent to say that "homological" is homological than to say that it is

---

<sup>25</sup> This paradox was inspired by Russell's paradox.

<sup>26</sup> Germany, 1886-1942. The paradox is also called the Grelling-Nelson paradox, because it was presented in a 1908 paper written jointly with Leonard Nelson (Germany, 1882-1927).

heterological, the sequence of predicates would seem consistent both ways, i.e.:

- If “homological” is homological, then it is homological (consistent predicates).
- If “homological” is heterological, then it is heterological (consistent predicates).

This could be taken to suggest that the term homological is somehow better constructed, while the term heterological has a structural fault. But this is not the real issue here.

The real issue is distinguishing between the physical words “homological” and “heterological” and their respective intended meanings, viz. homological and heterological. When we intend a word as such, we traditionally place it in inverted commas; and when we intend its assigned meaning we use it simply. In the above propositions, through which a paradox apparently arises, the subjects are words as such (in inverted commas) and the predicates are the meanings of such words.

In this perspective, there is no basis for the claim that “heterological” is heterological implies “heterological” is homological, or vice versa. The inference is very superficial, because it confuses the word as such (intended as the subject) with the meaning of the word (intended as the predicate). That is, the inverted commas in the subject are not used sincerely, but we secretly intend the underlying meaning as our subject.

*How did we draw out the consequents from the antecedents?* Could we see at a glance that the first thesis implies the second? Let us look at the hypothetical propositions in question more closely:

If in the antecedent we place the emphasis on *the property referred to* by the word “heterological”, viz. some presumed quality called heterologicality, we would formulate the paradoxes as follows:

- If the word “heterological” has the property it refers to (i.e. it is heterological), then it apparently lacks the property it refers to (i.e. is homological).
- If the word “heterological” lacks the property it refers to (i.e. it is homological), then it apparently has the property it refers to (i.e. is heterological).

If on the other hand, in the antecedent we place the emphasis on the word “heterological” *having or lacking* the property it refers to, we would instead formulate the paradoxes as follows:

- If the word “heterological” has the property it refers to (i.e. it is homological), then it apparently lacks the property it refers to (i.e. is heterological).
- If the word “heterological” lacks the property it refers to (i.e. it is heterological), then it apparently has the property it refers to (i.e. is homological).

In any of these cases, the consequent is constructed by comparing the subject “heterological” to the antecedent predicate heterological or homological; if they are the same word, we ‘infer’ homological as our consequent predicate, while if they verbally differ, we ‘infer’ heterological. But in truth, in making these comparisons between antecedent subject and predicate, we have not spotted any quality in the word “heterological” as such, but have tacitly referred to its underlying meaning, and faced that off against the hypothesized predicate.

In other words, the statement that “heterological” is homological (or for that matter that “homological” is heterological) is not as self-contradictory as it appears at first glance; it could conceivably be consistent. In truth, it is indeterminate and therefore meaningless.

More precisely, to resolve the paradox we have to remember how our terms were induced in the first place. We can tell that “short” is short *merely by seeing or hearing*

*the word* “short” (supposing that any one syllable, however written or pronounced, counts as short). But in the case of a term like heterological, you cannot tell whether the word has or lacks the property it refers to, because that property is not a concrete (visible or audible) quality of the word, but something abstract that we apply to visible or audible components of words. If the quality sought is not visible or audible, it is unknowable and there is no way for us to tell which predicate applies.

That is, our initial definitions of those terms, which mention “a word having/lacking a certain quality it refers to”, are not clear and precise, because they do not specify as they should that the qualities intended are phenomenal, i.e. perceptible aspects of the word. If the word labels something not included in its physical aspects, the terms homological and heterological simply do not apply. To apply them is to play verbal tricks. Thus, neither of these predicates is applicable to either of these words as such.

It might be objected that words do have non-phenomenal attributes. For example, we often consider a word useful or useless. In such case, we might ask: is the word “useful” useful or not? Yes, I’d reply to that. Therefore, “useful” is homological. Likewise, “useless” is useful, therefore “useless” is heterological. In this perspective, one may doubt the exactitude of what we have just proposed, that homological and heterological are terms that presuppose concrete (rather than abstract) predicates.

But to this objection, one could counter that the utility of a word is ultimately something concrete: a word is useful if it makes a perceptible practical difference in the development of knowledge. In that case, our definition could be modified slightly, specifying that the terms homological or heterological are only applicable when we can first *directly or indirectly* anchor them to some concrete property.

In sum, *these terms must refer to something other than themselves before they can at all be used*. The fallacy involved is similar to that in the liar paradox, where the term “this” is used with reference to itself, whereas it only acquires meaning when it has something else to refer to. Such terms are *relational*, and so cannot refer to other relations in a circular manner or ad infinitum: they need to eventually be anchored to some non-relational term.

Notice, by the way, that if we changed the word “short” to say “shortissimo”, with reference to the same meaning, the word would change status and become heterological, since “shortissimo” is not shortissimo. On the other hand, whatever other word we substitute for the word “heterological”, Grelling’s paradox in relation to it remains apparent. This test shows that in the latter case it is not purely the word that we are thinking of, but rather its underlying meaning. With regard to the word “useful”, we could also say that it is useful by virtue of its content, or at most by virtue of its being a word (a unit of language), and not because of its specific shape or sound.

## 8. The Russell paradox (redux)

*From A Fortiori Logic, appendix 7.5.*

Logic is what helps us transmute scattered concrete perceptions into well-ordered abstract concepts. Human knowledge, or opinion, is based on experience, imagination and rational insight. The latter is a kind of ‘experience’ in the larger sense, a non-phenomenal sort of experience, call it logical ‘intuition’. Reason was for this reason called by the ancients, in both West and East, the ‘sixth sense’ or ‘common sense’, i.e. the sense-organ which ties together the other five senses, those that bring us in empirical contact with phenomenal experience: colors, shapes, sounds, smells, tastes, touch-sensations, etc., whether they are physically perceived or mentally imagined. The five senses without the sixth yield chaotic nonsense (they are non-sense, one cannot ‘make sense’ of them); and conversely, the sixth sense is useless without the other five, because it has nothing about which to have rational insights. Imagination reshuffles past experiential data and reasoning, making possible the formation of new ideas and theories which are later tested with reference to further experience and reasoning.

### 1. Elements of class logic

Logic initially developed as a science primarily with reference to natural discourse, resulting in what we today refer to as predicate logic. In natural human discourse, we (you and me, and everyone else) routinely think of and discuss things we have perceived, or eventually conceived,

by means of categorical propositions involving a subject (say, S) and a predicate (say, P) which are related to each other by means of the copula 'is'. Such propositions have the form "S is P," which may be singular or plural, and in the latter case general (or universal) or particular, and positive or negative, and moreover may involve various modes and categories of modality<sup>27</sup>.

A proposition of the form 'S is P' is really a double predication – it tells us that a thing which is S is also P; thus, S and P are really both predicates, though one (the subject S) is given precedence in thought so as to 'predicate' the other (the predicate P) of it<sup>28</sup>. Primarily, S refers to some concrete phenomenon or phenomena (be it/they physical, mental or spiritual), i.e. an individual entity or a set of entities, and P to a property of it or of theirs. For examples, "John is a man" and "All men are human beings" are respectively a singular predication (about one man, John) and a plural one (about all men).

Additionally, still in natural discourse, the subject of our thoughts may be predicates *as such*, i.e. predicates in their capacity as predicates; an example is: "'men' may be the subject or predicate of a proposition." The latter occurs in specifically philosophical (or logical or linguistic) discourse; for example, in the present essay.

Now, logicians through the ages, and especially in modern times, have effectively found natural discourse somewhat inadequate for their needs and gradually developed a more

---

<sup>27</sup> We need not go into the details of these distinctions here, for they are well known. There are also many fine distinctions between different sorts of terms that may appear in propositions as subjects or predicates; but let us keep the matter simple.

<sup>28</sup> 'Predication' refers to the copula and the predicate together as if they were an action of the speaker (or the statement made) on the subject.



artificial language, that of ‘classes’<sup>29</sup>. This type of discourse exactly parallels natural discourse, but is a bit more abstract and descriptive so as to facilitate philosophical (or logical or linguistic) discourse and make it more precise. In this language, instead of saying “this S is P,” we say “this S is a member (or instance) of P” (note well the lengthening of the copula from ‘is’ to ‘is a member (or instance) of’). If ‘this S’ symbolizes a concrete individual, then ‘P’ here is called a ‘class’; but if ‘this S’ symbolizes an abstract class, then ‘P’ here is called a ‘class of classes’.

A class, then, is an abstraction, a mental construct in which we figuratively group some concrete things (be they physical, mental or spiritual). Although we can and do temporarily mentally classify things without naming the class for them, we normally name classes (i.e. assign them a distinctive word or phrase) because this facilitates memory and communication. Naming is not the essence of classification, but it is a great facilitator of large-scale classification. The name of a class of things does not ‘stand for them’ in the way of a token, but rather ‘points the mind to them’ or ‘draws our attention to them’; that is to say, it is an instrument of intention.

A class in the primary sense is a class of things in general; a class in the secondary sense is more specifically a class of classes. Membership is thus of two kinds: membership of non-classes in a class, or membership of classes in a class of classes. Alternatively, we may speak of first-order

---

<sup>29</sup> The following account of class logic is based on my presentation in *Future Logic*, chapters 43-45. The word ‘class’ comes from the Latin *classis*, which refers to a “group called to military service” (Merriam-Webster). I do not know whether the Ancients used that word in its logical sense, or some such word, in their discourse, but they certainly thought in class logic mode. Examples of class thinking are Aristotle’s distinction between species and genera and Porphyry’s tree.

classes and second-order classes to distinguish these two types. There are no other orders of classes. When we think about or discuss more concrete things, we are talking in first-order class-logic; when we think about or discuss first-order classes, we are talking in second-order class-logic, and the latter also applies to second-order classes since after all they are classes too. The two orders of classes should not be confused with the hierarchy of classes within each order.

The relation between classes of classes and classes is analogous to the relation between classes and concretes; it is a relation of subsumption. When a lower (i.e. first-order) class is a member of a higher (i.e. second-order) class, it does *not* follow that the members of the lower class are also members of the higher class; in fact, if they are members of the one they are certainly not members of the other. Thus, for example, you and me, although we are members of the class ‘men’ because we are men, we are not members of the class ‘classes of men’ because we are not ‘men’. Also, the class ‘men’ is not a man, but is a member of the class ‘classes of men’. The members of the class ‘classes of men’ (or more briefly put, ‘men-classes’), which is a class of classes, are, in addition to the broad class ‘men’, the narrower classes ‘gardeners’, ‘engineers’, ‘sages’, ‘neurotics’, and so on.<sup>30</sup>

Hierarchization, on the other hand, refers to classes *within a given order* that share instances, not merely by partly

---

<sup>30</sup> Note that saying or writing the word men without inverted commas refers to a *predicate*. When we wish to refer to the corresponding *class*, we say the class of men, or the class men; if we are writing, we may write the same with or without inverted commas, or simply ‘men’ in inverted commas. When dealing with classes of classes, we say the class of classes of men, or the class of men-classes, or the class men-classes, and we may write the same with or without inverted commas, or simply ‘classes of men’ or ‘men-classes’ in inverted commas.

overlapping, but in such a way that all the members of one class are members of the other (and in some but not all cases, vice versa). For example, since all men are animals, though not all animals are men, the class 'men' is a *subclass (or species)* of the class 'animals', and the class 'animals' is an *overclass (or genus)* of the class 'men'. If two classes have the same instances, no more and no less, they may be said to be *co-extensive* classes (a class that serves as both species and genus in some context is said to be *sui generis*). If two classes merely share some instances, they may be said to be *intersecting (or overlapping)* classes, but they are not hierarchically arranged (e.g. 'gardeners' and 'engineers'). If two classes of the same order have no instances in common, they may be said to be *mutually exclusive* classes.

It is important to grasp and keep in mind the distinction between hierarchy and order. Since you and I are men, each of us is a member of the class 'men'; this is subsumption by a first-order class of its concrete instances. Since all men are animals, the class 'men' is a subclass of the class 'animals'; this is hierarchy between two classes of the first order. Since 'men' is a class of animals, it is a member of the class 'classes of animals' (or 'animal-classes'); this is subsumption by a second-order class (i.e. a class of classes) of its first-order-class instances (i.e. mere classes). Since all 'classes of men' are 'classes of animals', the class 'men-classes' is a subclass of the class 'animals-classes'; this is hierarchy between two classes of the second order, i.e. between two classes of classes. The relation between classes of the first order and classes of the second order is never one of hierarchy, but always one of subsumption; i.e. the former are always members (instances) of the latter, never subclasses. Hierarchies only occur between classes of the same order.

Thus, in class logic, we have two planes of existence to consider. At the ground level is the relatively objective

plane of empirical phenomena (whether these are physical, mental or spiritual in substance); above that, residing in our minds, is the relatively subjective plane of ideas (which are conceived as insubstantial, but do have phenomenal aspects – namely mental or physical images, spoken or written words, and the intentions of such signs), comprising ideas about empirical phenomena and ideas about such ideas. Classes are developed to facilitate our study of empirical phenomena and classes of classes are developed in turn to facilitate our study of classes – for classes (including classes of classes) are of course themselves empirical phenomena of sorts. Classification is a human invention helpful for cognitive ordering of the things observed through our senses or our imaginations or our introspective intuitions. Although classes are products of mind, this does not mean that they are arbitrary – they are formed, organized and controlled by means of our rational faculty, i.e. with the aid of logic.

Clearly, to qualify as a class, a class must have at least one member (in which case the sole member is “one of a kind”). Usually, a class has two or more members, indeed innumerable members. A class is finite if it includes a specified number of instances; if the number of instances it includes is difficult to enumerate, the class is said to be open-ended (meaning infinite or at least indefinite). What brings the instances of a class together in it is their possession of some distinctive property in common; the class is defined by this property (which may of course be a complicated conjunction of many properties). A class without instances is called a null (or empty) class; this signifies that its defining property is known to be fanciful, so that it is strictly speaking a non-class.

Thus, note well, the term ‘class’ is a bit ambiguous, as it may refer to a first-order class (a class of non-classes, i.e. of things other than classes) or a second-order class (a class of classes, i.e. a mental construct grouping two or more

such mental constructs). A class (of the first order) is not, indeed cannot be, a class of classes (i.e. a class of the second order). There is, of course, a class called 'non-classes'; its instances are principally all concrete things, which are not themselves classes; for example, you and I are non-classes. 'Non-classes' is merely a class, not a class of classes, since it does not include any classes. Thus, 'non-classes' may be said to be a first-order class, but does not qualify as a second-order class.<sup>31</sup>

The realm of classes of classes is very limited as an object of study in comparison to the realm of mere classes. For what distinctions can we draw between classes? Not many. We can distinguish between classes and classes of classes, between finite and open-ended classes, between positive and negative classes<sup>32</sup>, and maybe a few more things, but not much more.

## 2. An apparent double paradox

Bertrand Russell (Britain, 1872-1970) proposed a distinction between 'a class that is a member of itself' and

---

<sup>31</sup> Note that, whereas positive terms are easy enough to translate into class logic language, negative terms present a real difficulty. For example, whereas the term men refers only to non-classes, its strict antithesis, the term non-men in its broadest sense, includes both non-classes (i.e. concrete things other than men) and classes (i.e. more abstract things). Again, whereas the term finite classes refers only to classes, its strict antithesis, the term non-finite-classes in its broadest sense, includes both open-ended classes (abstracts) and non-classes (concretes). Thus, we must, for purposes of consistency, admit that some terms do cover both non-classes and classes (including classes of classes). Practically, this means we have to make use of *disjunctives* which reveal the implicit alternatives. This of course complicates class logic considerably.

<sup>32</sup> Positive classes are defined by some positive property and negative classes are defined by a negative one. For examples, 'men' is defined with reference to rational animals (positive), whereas 'bachelors' is defined with reference to not yet married men (negative).

‘a class that is not a member of itself’. Although every class is necessarily co-extensive with itself (and in this sense is included in itself), it does not follow that every class is a member of itself (evidently, some are and some are not). Such a distinction can be shown to be legitimate by citing convincing examples. Thus, the class ‘positive classes’ is a member of itself, since it is defined by a positive property; whereas the class ‘negative classes’ is not a member of itself, since it is also positively defined (albeit with general reference to negation). Again, the class ‘finite classes’ is not a member of itself, since it has innumerable members; while the class ‘open-ended classes’ is a member of itself, since it too has innumerable members.

What about the class ‘classes’ – is it a member of itself or not? Since ‘classes’ is a class, it must be a member of ‘classes’ – i.e. of itself. This is said without paying attention to the distinction between classes of the first and second orders. If we ask the question more specifically, the answer has to be nuanced. The class ‘first-order classes’ being a class of classes and not a mere class, cannot be a member of itself, but only a member of ‘second-order classes’; the members of the ‘first-order classes’ are all mere classes. On the other hand, since the class ‘second-order classes’ is a class of classes, it is a member of itself, i.e. a member of ‘second-order classes’. Thus, the class ‘second-order classes’ includes both itself and the class ‘first-order classes’, so that when we say that the wider class ‘classes’ is a member of itself, we mean that it is more specifically a member of the narrower class ‘classes of classes’. As regards the class ‘non-classes’, since it is a class and not a non-class, it is not a member of itself. Note however that Russell’s paradox does not make a distinction between classes of the first and second orders, but focuses on ‘classes’ indiscriminately.

Russell asked **whether “the class of all classes which are not members of themselves” is or is not a member of**

**itself.** It seemed logically impossible to answer the question, because either way a contradiction ensued. For if the class ‘classes not members of themselves’ *is not* a member of the class ‘classes not members of themselves,’ then it is indeed a member of ‘classes not members of themselves’ (i.e. of itself); and if the class ‘classes not members of themselves’ *is* a member of ‘classes not members of themselves,’ then it is also a member of ‘classes which are members of themselves’ (i.e. of its contradictory). This looked like a mind-blowing double paradox.

**The solution of the problem.** The pursuit of knowledge is a human enterprise, and therefore one which proceeds by trial and error. Knowledge is inductive much more than deductive; deduction is just one of the tools of induction. There are absolutes in human knowledge, but they are few and far between. When we formulate a theory, it is always essentially a hypothesis, which might later need to be revised or ruled out. So long as it looks useful and sound, and does so more than any competing theory, we adopt it; but if it ever turns out to be belied by some facts or productive of antinomy, we are obliged to either reformulate it or drop it. This is the principle of induction. When we come upon a contradiction, we have to ‘check our premises’ and modify them as necessary. In the case at hand, since our conception of class logic is shown by the Russell paradox to be faulty somehow, we must go back and find out just where we went wrong. So, let us carefully retrace our steps. We defined a class and membership in a class by turning predication into classification, saying effectively:

If something is X, then it is a member of the class ‘X’, and not a member of the class ‘nonX’.

If something is not X, then it is not a member of the class ‘X’, but a member of the class ‘nonX’.

Where did we get this definition? It is not an absolute that was somehow cognitively imposed on us. We invented it – it was a convention by means of which we devised the idea of classes and membership in them. Knowledge can very well proceed without recourse to this idea, and has done so for millennia and continues to do so in many people’s mind. It is an idea with a history, which was added to the arsenal of reasoning techniques by logicians of relatively recent times. These logicians noticed themselves and others reasoning by means of classification, and they realized that this is a useful artifice, distinct from predication and yet based on it somehow. They therefore formally proposed the above definition, and proceeded to study the matter in more detail so as to maximize its utility. The ‘logic of classes’, or ‘class logic’, was born.

However, at some stage, one logician, Bertrand Russell, realized that there was an inherent inconsistency in our conception of classification, which put the whole edifice of class logic in serious doubt. That was the discovery of the paradox bearing his name. That was a great finding, for there is nothing more important to knowledge development, and especially to development of the branch of knowledge called formal logic, than the maintenance of consistency. Every discovery of inconsistency is a stimulation to refine and perfect our knowledge. Russell deserves much credit for this finding, even if he had a lot of difficulty resolving the paradox in a fully convincing manner. Let us here try to do better, by digging deeper into the thought processes involved in classification than he did. What is classification, more precisely?

If we look more closely at our above definition of a class ‘X’ and membership of things in it by virtue of being X, we must ask the question: what does this definition achieve, concretely? Are we merely substituting the phrase ‘is a member of’ for the copula ‘is’, and the class ‘X’ for the predicate X? If this is what we are doing, there is no



point in it – for it is obvious that changing *the name* of a relation or a term in no way affects it. Words are incidental to knowledge; what matters is their underlying intent, their meaning. If the words change, but not the meaning, nothing of great significance has changed. No, we are not here merely changing the words used – we are proposing a mental image.

Our idea of classification is that of *mental entities* called classes in which things other than classes (or lesser classes, in the case of classes of classes) are *figuratively collected and contained*. When we say of things that they are members of class ‘X’, we mean that class ‘X’ is a sort of box *into* which these things are, by means of imagination, stored (at a given time, whether temporarily or permanently). That is to say, our ‘definition’ of classification is really a formal convention used to institute this image. What it really means is the following:

If something is X, then it is *in* the class ‘X’, and out of (i.e. *not in*) the class ‘nonX’.

If something is not X, then it is out of (i.e. *not in*) the class ‘X’, but in the class ‘nonX’.

Clearly, to ‘be’ something and to ‘be in’ (within, inside) something else *are not the same thing*. Our definition conventionally (i.e. by common agreement) decrees that if X is predicated of something, then we may think of that thing as being as if contained by the mental entity called class ‘X’. But this decree is not an absolute; it is not a proposition that being subject to predication of X *naturally and necessarily implies* being a member of class ‘X’. For the whole idea of classification, and therefore this definition of what constitutes a class and membership therein, is a human invention. This invention may well be, and indeed is, very useful – but it remains bound by the laws of nature. If we find that the way we have conceived it, i.e. our definition of it, inevitably leads to contradiction,

we must adjust our definition of it in such a way that such contradiction can no longer arise. This is our way of reasoning and acting in all similar situations.

As we shall presently show, since the contradiction is a consequence of the just mentioned defining implication, we must modify that implication. That is to say, we must decree it to have limits. Of course, we cannot just vaguely say that it has limits; we must precisely define these limits so that the practical value of our concept of classification is restored. We can do that by realizing that our definition of classification with reference to something ‘being in’ something else means that *class logic is conceived of as related to geometrical logic*. This is obvious, when we reflect on the fact that we often ‘represent’ classes as geometrical figures (notably, circles) and their members as points within those figures. This practice is not accidental, but of the very essence of our idea of classification. Classification is imagining that we put certain items, identified by their possession of some common and distinctive property, in a labeled container<sup>33</sup>.

Let us now examine the concept of self-membership in the light of these reflections. What is the idea of self-membership? It is the presupposition that a class may be a member of itself. But is that notion truly conceivable? If we for a moment put aside the class logic issue, and reformulate the question in terms of geometrical logic, we see that it is absurd. Can a container contain itself? Of course not. There is no known example of a container containing itself in the physical world; and indeed, we cannot even visually imagine a container containing itself. So, the idea of self-containment has no empirical basis, not even in the mental sphere. It is only a fanciful conjunction of two words, without experiential basis. For this reason,

---

<sup>33</sup> This is a pictorial ‘representation’, an analogical image not to be taken literally.

the idea strikes us as illogical and we can safely posit as a universal and eternal ‘axiom’ that self-containment is impossible. A nonsensical term like ‘the collection of all collections’ is of necessity an empty term; we are not forced to accept it, indeed we are logically not allowed to do so; we can only consistently speak of ‘the collection of all *other* collections’<sup>34</sup>.

A container is of course always co-extensive with itself, i.e. it occupies exactly the space it occupies. But such ‘co-extension’ is not containment, let alone self-containment, for it does not really (other than verbally) concern two things but only one; there is no ‘co-’ about it, it is just extended, just once. We refer to containment when a smaller object fits inside a larger object (or in the limit when *another* object of equal size neatly fits inside a certain object). The concept of containment refers to two objects, not one. There has to be two distinct objects; it does not suffice to label the same object in two ways. To imagine ‘self-containment’ is to imagine that a whole object can somehow fit into itself as a smaller object (or that it can somehow become two, with one of the two inside the other). This is *unconscionable*. A whole thing cannot be a part (whether a full or partial part) of itself; nothing can be both whole and part at once. A single thing cannot be two things (whether of the same or different size) at once; nothing can simultaneously exist as two things.

You cannot decide by convention that something is both whole and part or that one thing is two. You cannot convene something naturally impossible. You can only convene something naturally possible, even though it is unnecessary. Thus, the concept of self-containment is

---

<sup>34</sup> To give a concrete image: a bag of marbles (whether alone or, even worse, with the marbles in it) cannot be put inside itself, even if the bag as a whole, together with all its contents, can be rolled around like a marble and so be called a marble.

meaningless; it is an inevitably empty concept, because it assumes something impossible to be possible. There is no such thing as self-containment; a container can never contain itself. If this is true, then it is of course equally true that no class includes itself, for (as we have seen) classification is essentially a geometrical idea. Given that a container cannot contain itself, it follows that the answer to the question as to whether a class can be a member of itself is indubitably and definitely: No. Because to say of any class that it is a member of itself is to imply that a container can be a content of itself. Just as no container which is a content of itself exists, so *no class which is a member of itself exists!*

Now, this is a revolutionary idea for class logic. It applies to any and every class, not just to the class ‘classes not members of themselves’ which gave rise to the Russell paradox. Moreover, note well that we are here denying the possibility of membership of a class in itself, but not the possibility of *non*-membership of a class in itself. When we say that no container contains itself, we imply that it is true of each and every container that it does not contain itself. Similarly, when we say that no class is a member of itself, we imply that it is true of each and every class that it is not a member of itself. What this means is that while we acknowledge the subject of the Russell paradox, namely the class ‘classes that are not members of themselves’, we reject the notion that such a class might *ever*, even hypothetically for a moment, be a member of itself (and therefore paradoxical) – for, we claim, no class whatever is ever a member of itself.

How can this be, you may well ask? Have we not already shown by example that some classes are members of themselves? Have we not agreed, for example, that the class ‘classes’ being a class has to be a member of the class ‘classes’, i.e. of itself? How can we deny something so obvious? Surely, you may well object further, if the class

'classes that are not members of themselves' is not a member of itself, then it is undeniably a member of itself; and if it is a member of itself, then it is undeniably not a member of itself? To answer these legitimate questions, let us go back to our definition of classification, and the things we said about that definition. As I pointed earlier, our definition of classes and membership in them has the form of a conventional implication. It says:

If and only if something is X, then it is a member of the class 'X'.

Now, since this conventional implication leads us inexorably to paradox, we must revise it, i.e. make it more limited in scope, i.e. specify the exact conditions when it 'works' and when it ceases to 'work'. What is essentially wrong with it, as we have seen, is that it suggests that a class can be a member of itself. For example, since the class 'classes' is a class, then it is a member of 'classes'; in this example, the variable X has value class and the variable 'X' has value 'classes'. But, as we have shown, the claim that a class can be a member of itself logically implies something geometrically impossible; namely, that a container can be a content of itself. So, to prevent the Russell paradox from arising, we need to prevent the unwanted consequences of our definition from occurring. Given that our concept of classification is problematic as it stands, what are the conditions we have to specify to delimit it so that the problem is dissolved?

The answer to this question is that when the subject and predicate of the antecedent clause are one and the same, then the consequent clause should cease to be implied. That is to say, if the antecedent clause has the form "if the class 'X' is X" then the consequent clause "then the class 'X' is a member of 'X' (and thus of itself)" *does not follow*. This 'does not follow' is a convention, just as the general 'it follows' was a convention. What we have done here is

merely to draw a line, saying that the consequent *generally* follows the antecedent, *except in the special case* where the subject and predicate in the antecedent are 'the same' (in the sense that predicate X is applicable to class 'X' which is itself based on predicate X). This is logically a quite acceptable measure, clearly. If an induced general proposition is found to have exceptions, then it is quite legitimate and indeed obligatory to make it less general, retreating only just enough to allow for these exceptions.

Since the initial definition of classification was a general convention, it is quite permissible, upon discovering that this convention leads us into contradiction, to agree on a slightly narrower convention. Thus, whereas, in the large majority of cases, it remains true that if something is X, then it is a member of the class 'X', and more specifically, if a class (say, 'Y') is X, then it (i.e. 'Y') is a member of the class of classes 'X' – nevertheless, *exceptionally, in the special case where the class that is X is the class 'X' (i.e. where 'Y' = 'X'), we cannot go on to say of it that it is a member of 'X'*, for this would be to claim it to be a member of itself, which is impossible since this implies that a container can be a content of itself. Note well that we are not denying that, for example, the class 'classes' *is* a class; we are only denying *the implication* this is normally taken to have that the class 'classes' *is a member of* the class 'classes'. We can cheerfully continue saying 'is' (for that is mere predication), but we are not here allowed to turn that 'is' into 'is a member of' (for that would constitute illicit classification).

In this way, the Russell paradox is inhibited from arising. That is to say, with reference to the class 'classes not members of themselves': firstly, it is quite legitimate to suppose that the class 'classes not members of themselves' is not a member of itself, since we know for sure (from geometrical logic) that no class is a member of itself; but it is *not* legitimate to say that this fact (i.e. that it is not a

member of itself) implies that it is a member of itself, since such implication has been conventionally excluded. Secondly, it is *not* legitimate to suppose, even for the sake of argument, that the class ‘classes not members of themselves’ is a member of itself, since we already know (from geometrical logic) that no class is a member of itself, and therefore we cannot establish through such supposition that it is not a member of itself, even though it is anyway true that it is not a member of itself.

As can be seen, our correction of the definition of classification, making it less general than it originally was, by specifying the specific situation in which the implication involved is not to be applied, succeeds in eliminating the Russell paradox. We can say that the class ‘classes not members of themselves’ is not a member of itself, but we cannot say that it is a member of itself; therefore, both legs of the double paradox are blocked. In the first leg, we have blocked the inference from not-being ‘a member of itself’ to being one; in the second leg, we have interdicted the supposition of being ‘a member of itself’ even though inference from it of not-being one would be harmless. Accordingly, the answer to the question posed by Russell – viz. “Is the class of all classes which are not members of themselves a member of itself or not?” – is that this class is not a member of itself, and that this class not-being a member of itself does not, contrary to appearances, make it a member of itself, because no class is a member of itself anyway.

Thus, to be sure, though it is true that the class ‘classes’ is a class, it does not follow that it is a member of itself; though it is true that the class ‘classes of classes’ is a class of classes, it does not follow that it is a member of itself; though it is true that the class ‘positive classes’ is a positive class, it does not follow that it is a member of itself; though it is true that the class ‘open-ended classes’ is an open-ended class, it does not follow that it is a member of itself;

though it is true that the class ‘classes that are not members of themselves’ is a class that is not a member of itself, it does not follow that it is a member of itself. As for the class ‘classes members of themselves’, it has no members at all. It should be emphasized that the restriction on classification that we have here introduced is of very limited scope; it hardly affects class logic at all, concerning as it does a few very borderline cases.

The above is, I believe, the *correct and definitive resolution* of the Russell paradox. We acknowledged the existence of a problem, the Russell paradox. We diagnosed the cause of the problem, the assumption that *self-membership* is possible. We showed that self-membership is unconscionable, since it implies that a container can contain itself; this was not arbitrary tinkering, note well, but appealed to reason. We proposed a solution to the problem, one that precisely targets it and surgically removes it. Our remedy consisted in *uncoupling* predication from classification in all cases where self-membership is assumed, and only in such cases. This solution of the problem is plain common sense and not a flight of speculation; it is simple and elegant; it is convincing and uncomplicated; it does not essentially modify the concept of class membership, but only limits its application a little; it introduces a restriction, but one that is clearly circumscribed and quite small; it does not result in collateral damage on areas of class logic, or logic in general, that are not problematic, and therefore does not call for further adaptations of logic doctrine. Note moreover that our solution does not resort to any obscure ‘system’ of modern symbolic logic, but is entirely developed using ordinary language and widely known and accepted concepts and processes.



### 3. A bit of the history

Let us now look briefly at some of the history of the Russell paradox, and see how he and some other modern symbolic logicians dealt with it<sup>35</sup>.

Georg Cantor had already in 1895 found an antinomy in his own theory of sets. In 1902, when Gottlob Frege (Germany, 1848-1925) was about to publish the second volume of his *Grundgesetze*, he was advised by Russell of the said paradox. Frege was totally taken in and could not see how to get out of the self-contradictions inherent in “the class of classes that do not belong to themselves.” He perceived this as very serious, saying: “What is in question is ... whether arithmetic can possibly be given a logical foundation at all.” Frege first tried to fix things by suggesting that there might be “concepts with no corresponding classes,” or alternatively by adjusting one of his “axioms” in such a way that:

“Two concepts should be said to have the same extension if, and only if, every object which fell under the first *but was not itself the extension of the first* fell likewise under the second and vice versa”<sup>36</sup>.

Clearly, Frege’s initial suggestion that there might be “concepts with no corresponding classes” can be viewed as an anticipation of my uncoupling of predication and classification in specific cases. However, Frege did not identify precisely in what cases such uncoupling has to occur. This is evident in his next suggestion, which, though it points tantalizingly to the difficulty in the notion of self-

---

<sup>35</sup> I am here referring principally to the account by William and Martha Kneale in *The Development of Logic*, ch. XI.1-2.

<sup>36</sup> Kneale and Kneale, p. 654. Italics theirs.

membership, does not reject this notion outright but instead attempts to mitigate it. He speaks of two concepts instead of one, and tries to conventionally exclude the extension as a whole of each from the other, while of course continuing to include the objects falling under the extension; this shows he has not realized that self-inclusion by an extension is not even thinkable.

It should be stressed that Russell's paradox pertains to a certain class (viz. that of all classes not members of themselves) being or not-being a member of *itself* – not of *some other* class. Frege tries to resolve this paradox with reference, not to a single class, but to a pair of equal classes, even though (to my knowledge) he has not demonstrated that co-extensive classes result in a paradox comparable to the Russell paradox. It follows that his attempted solution to the problem is not germane to it. Moreover, Frege seems to have thought that if all items that fall under one class (say, 'Y') fall under another class (say, 'X'), then *the class* 'Y' may be assumed to fall under the class 'X'; and vice versa in the event of co-extension. This is suggested by his attempt to prevent such assumption, so as to avoid (in his estimate) the resulting Russell paradox. But in truth, it does not follow from the given that all Y are X that the class 'Y' is a member of the class 'X' – it only follows that the class 'Y' is a subclass of the class 'X', or an equal class if the relation is reversible. Thus, it appears that Frege confused the relations of class-membership and hierarchization of classes, using a vague term like 'falling' to characterize them both.

We may well ask the question whether an equal class, or a subclass, or even an overclass, might be a member of its hierarchically related class. Offhand, it would seem to be possible. For example, all positive classes are classes and therefore members of the class 'classes', and the class 'positive classes' is a subclass of the class 'classes'; however, although not all classes are positive classes (some

are negative classes), nevertheless the class ‘classes’ is a positive class (being defined by a positive statement), and so is a member of the class ‘positive classes’. But although this example suggests that an overclass might be a member of its subclass (and therefore, all the more, an equal class or a subclass might be a member of its hierarchical relative), we might still express a doubt by means of *analogy*, as Frege perhaps intended to do. We could argue inductively, by generalization, that if a class cannot be a member of itself, then maybe it cannot be a member of any coextensive class (as Frege suggests), and perhaps even of a subclass or an overclass. For the issue here is whether the instances referred to by the first class can be thought to occur twice in the second class (as members of it in their own right, and as constituents of a member). So, Frege may have raised a valid issue, which could lead to further restrictions in class logic. However, this need not concern us further in the present context, since (as already explained) it is not directly relevant to resolution of the Russell paradox.

Russell described his paradox in his book *Principles of Mathematics*, published soon after. Although at first inclined to Frege’s second approach, he later preferred Frege’s first one, proposing that there might be “some propositional functions which did not determine genuine classes.” Note here again the failure to pinpoint the precise source and remedy of the problem. Subsequently, Russell thought that “the problem could never be solved completely until all classes were eliminated from logical theory.” This, in my view, would be throwing out the baby with the bath water – an overreaction. But then he found out (or rather, he thought he did, or he convinced himself that he did) that the same paradox could be generated without “talk of classes,” i.e. with reference to mere

predicates – that is, in terms of predicate logic instead of in terms of class logic. As Kneale and Kneale put it<sup>37</sup>:

“Instead of the class which is supposed to contain all classes that are not members of themselves let us consider the property of being a property which does not exemplify itself. If this property exemplifies itself, then it cannot exemplify itself; and if it does not exemplify itself, then it must exemplify itself. Clearly, the nature of the trouble is the same here as in the original paradox, and yet there is no talk of classes.”

But even if classes are not explicitly mentioned here, it is clear that they are tacitly intended. How would a property ‘exemplify’ itself? Presumably, property X would be ‘a property which exemplifies itself’ if property X happens to be one of the things that have property X. That is to say, X exemplifies X if X is a member of the class of things that are X. We cannot talk about properties without resorting to predication; and once we predicate we can (given the initial definition of classification) surely classify. So, this attempt is just verbal chicanery; the same thought is intended, but it is dressed up in other words. It is dishonest. Moreover, the way the paradox is allegedly evoked here does not in fact result in paradox.

We cannot say, even hypothetically, “if this property [i.e. the property of being a property which does not exemplify itself] exemplifies itself” for that is already self-contradictory. To reconstruct a Russell paradox in ‘property’ terms, we would have to speak of ‘*the property of all properties* which do not exemplify themselves’; for then we would have a new term to chew on, as we did in

---

<sup>37</sup> William and Martha Kneale, p. 655.

class logic. But clearly, this new term is quite contrived and meaningless. Here again, we must mean ‘*the class of all properties which do not exemplify themselves*’ – and in that event, we are back in class logic. Thus, note well, while Russell was right in looking to see whether his paradox was a problem specific to class logic, or one also occurring in predicate logic, and he claimed to have established that it occurred in both fields, in truth (as we have just demonstrated) he did not succeed in doing that. In truth, the paradox *was* specific to class logic; and he would have been better off admitting the fact than trying to ignore it.

In response to certain criticisms by his peers, Russell eventually “agreed that the paradoxes were all due to vicious circles, and laid it down as a principle for the avoidance of such circles that ‘whatever involves all of a collection must not be one of the collection’.” Thus, Russell may be said to have conceded the principle I have also used, namely that a collection cannot include itself as one of the items collected, although in truth the way he put it suggests he conceived it as a convention designed to block incomprehensible vicious circles rather than a logical absolute (notice that he says ‘must not’ rather than ‘cannot’). He viewed the paradoxes of set theory as “essentially of the same kind as the old paradox of Epimenides (or the Liar).” This suggests that, at this stage, he saw his own paradox as due to self-reference, somehow. It does look at first sight as if there is some sort of self-reference in the proposition ‘the class of all classes that are not members of themselves is (or is not) a member of itself’, because the clause ‘member of itself’ is repeated (positively or negatively, in the singular or plural) in subject and predicate<sup>38</sup>. But it cannot be said that self-reference is exactly the problem.

---

<sup>38</sup> Note that if self-reference were the crux of the problem, then the proposition ‘the class of all classes that are members of themselves

A few years later, in a paper published in 1908, Russell came up with a more elaborate explanation of the Russell paradox based on his ‘theory of types’. Russell now argued that “no function can have among its values anything which presupposes the function, for if it had, we could not regard the objects ambiguously denoted by the function as definite until the function was definite, while conversely ... the function cannot be definite until its values are definite”<sup>39</sup>. In other words, the question “the class of all classes that are not members of themselves, is it or is it not a member of itself” is inherently flawed, because the subject remains forever out of reach. We cannot take hold of it till we resolve whether or not it is a member of itself, and we cannot do the latter till we do the former; so, the conundrum is unresolvable, i.e. the question is unanswerable. Effectively, the subject is a term cognitively impossible to formulate, due to the double bind the issue of its definition involves for any thinker.

Here, we should note that the purpose of Russell’s said explanation was effectively to invalidate the negative class ‘classes not members of themselves’, since this is the class giving rise to the double paradox he was trying to cure. The positive class ‘classes members of themselves’ clearly does not result in a double paradox: if we suppose it is *not* a member of itself, self-contradiction does ensue, but we can still say without self-contradiction that it *is* a member of itself. In fact, if Russell’s explanation were correct, the positive class ought to be as illicit as the negative one. For if we claim the impossibility of a class referring to something that is *not yet settled*, as Russell did with reference to the negative class, then we must admit this characteristic is also found in the positive class, and we

---

is (or is not) a member of itself’ would be equally problematic, even though it apparently does not result in a similar paradox.

<sup>39</sup> Quoted by the Kneales, p. 658.

must reject it too. Russell does not seem to have realized that, i.e. that his remedy did not technically differentiate the two classes and so could be applied to both. For this reason, his attempt to solve the Russell paradox with reference to circularity or infinity must be judged as a failure. In my own theory, on the other hand, it is the positive class (that of self-membership) which is invalid (and empty), since it is geometrically unthinkable, while the negative class (that of non-self-membership) remains quite legitimate (and instantiated), as indeed we would expect on the principle that all claims (including that of self-membership) ought to be deniable.

Anyway, Russell concluded, briefly put, that a function could not be a value of itself; and proposed that function and value be differentiated as two 'types' that could not be mixed together indiscriminately. But this theory is, I would say, too general, and it complicates matters considerably. As we have seen, we cannot refuse to admit that, for instance, 'classes' *is* a class; the most we can do is to deny that this implies that 'classes' *is a member of* itself. This is a denial of self-membership, not of self-predication or of self-reference. As regards the notion of 'types' and later that of 'orders within types', these should not be confused with the more traditional ideas of hierarchies and orders of classes, which we laid out earlier in the present essay. In truth, the resemblance between Russell's concepts and the latter concepts gives Russell's theory a semblance of credibility; but this appearance is quite illusory – these are very different sets of concepts. Russell's notion of 'types' is highly speculative and far from commonsense; while it might appear to solve the Russell paradox, it has ramifications that range far beyond it and incidentally invalidate traditional ideas that do not seem problematic<sup>40</sup>.

---

<sup>40</sup> See for a start the Kneales' critique of the 'theory of types' in ch. XI.2.

In short, it is a rough-and-ready, makeshift measure, and not a very convincing one.

Every paradox we come across is, of course, a signal to us that we are going astray somehow. Accordingly, the Russell paradox may be said to have been a signal to Frege, Russell, and other modern logicians, that something was wrong in their outlook. They struggled hard to find the source of the problem, but apparently could not exactly pinpoint its location. All the intricacy and complexity of their symbolic and axiomatic approach to logic could not help them, but rather obscured the solution of the problem for them. This shows that before any attempt at symbolization and axiomatization it is essential for logicians to fully understand the subject at hand in ordinary language terms and by means of commonsense. To my knowledge, the solution of the problem proposed in the present essay is original, i.e. not to be found elsewhere. If that is true, then the theory of class logic developed by modern symbolic logicians, which is still the core of what is being taught in universities today, needs to be thoroughly reviewed and revised.

#### **4. A bit of self-criticism**

As regards the resolution of the Russell paradox that I proposed over two decades ago in my *Future Logic*, the following needs to be said here. While I stand, in the main, by my theory of the logic of classes there (in chapters 43-44), I must now distance myself somewhat from my attempted resolution of the Russell paradox there (in chapter 45).

I did, to my credit, in that past work express great skepticism with regard to the notion of self-membership; but I did not manage to totally rule it out. I did declare: “Intuitively, to me at least, the suggestion that something can be both container and contained is hard to swallow,”



and I even postulated, in the way of a generalization from a number of cases examined, that “no class of anything, or class of classes of anything, is ever a member of itself,” with the possible exception of “things” or “things-classes” (although it might be said of these classes that they are not members of any classes, let alone themselves<sup>41</sup>); but still, I did not reject self-membership on principle, and use that rejection to explain and resolve the Russell paradox, as I do in the present essay.

This is evident, for instance, in my accepting the idea that “‘self-member classes’ is a member of itself.” The reason I did so was the thought that “whether self-membership is possible or not, is not the issue.” Superficially, this is of course true – the Russell paradox concerns the ‘class of all classes that are not members of themselves’, and not ‘the class of all classes that are members of themselves’. But in fact, as I have shown today, this is not true; acceptance of self-membership is the true cause of the Russell paradox, and non-self-membership is not in itself problematic.

Anyway, not having duly ruled out self-membership, I resorted to the only solution of the problem that looked promising to me at the time – namely, rejection of ‘permutation’ from “is (or is not) a member of itself” to “is (or is not) {a member of itself}” (notice the addition of curly brackets). That is to say, I proposed the logical interdiction of changing the *relation* of self-membership or non-self-membership into a predicable *term*. Now I see that this was wrong – it was an action taken *too late* in the process of thought leading up to the Russell paradox. It was a superficial attempt, treating a symptom instead of the disease. I did that, of course, because I thought this was “of

---

<sup>41</sup> Note that in this context I come up with the idea that the definition of membership might occasionally fail. But I did not at the time pursue that idea further, because I did not then analyze what such failure would formally imply.

all the processes used in developing these arguments, [the] only one of uncertain (unestablished) validity.” But in truth, it was not the only possible cause of the effect – there was a process *before* that, one of deeper significance, namely the transition from ‘is’ to ‘is a member of’. I did not at the time notice this earlier process, let alone realize its vulnerability; and for that reason, I did not attack it.

Clearly, I was on the right track, in that I sought for a place along the thought process at which to block development of the Russell paradox. But my error was to pick a place too late along that process. In fact, the right place is earlier on, as advocated in the present essay. The Russell paradox does not arise due to an illicit permutation, but due to the illicit transformation of a predicate into a class in cases where a claim of self-membership would ensue. And while the remedy proposed is even now in a sense ‘conventional’, the flaw it is designed to fix is quite real – it is that self-membership is in fact impossible and therefore can never be assumed true. My previous proposed solution to the problem only prevented the Russell paradox; it did not prevent self-membership, which is the real cause of the paradox. Thus, the solution I propose in the present essay is more profound and more accurate.

## 9. More on the Russell paradox

*From Topics (published only online so far).*

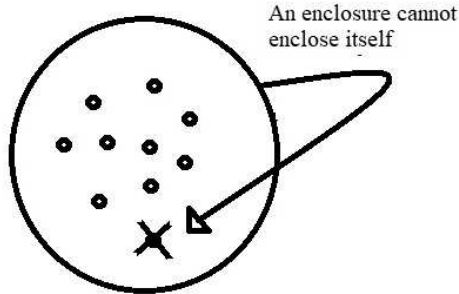
### 1. My resolution of the Russell paradox

My resolution of the Russell paradox is, simply put, that *no* class is, or can ever be, a member of itself. Self-membership is *unconscionable*. It is therefore no surprise that the unthinking adoption of the idea of self-membership by Russell led him to paradox. The concept of class-membership differs significantly from that of predication. Predication is a natural way of thinking; class-membership is an artificial one, invented by logicians. The concept of class-membership is useful, but it has limitations. The idea of class-membership is very early, being implied already in Aristotle's discourses; but became more consciously used by logicians in modern times.

The original idea is that if any individual thing (call it X) 'has' some property (call it Y-ness), then it is 'subsumed' under the species or genus of Y-having things; i.e. ***if X is Y, then X is a Y*** (notice the use of 'a' to signal that a species or genus, as distinct from a property, is intended by Y). This thought later led to the idea of a 'class with members', defined by: ***if X is a Y, then X is a member of the class "Y"*** (notice the use of inverted commas to refer to a class as distinct from a predicate). Tacitly intended here are the complements: "if X is not Y, then X is not a Y" and "if X is not a Y, then X is not a member of the class 'Y'," note.

But it should be made clear that these two ideas, though related, are quite distinct. While it can be said of anything

which is Y, that it is *a* Y; it cannot be said of any class “Y” which happens to be a Y, that it is a member of the class “Y” (i.e. a member of itself). For the idea of class-membership is figuratively the idea of an enclosure (the class) in which a group of things (the members) that are alike in some way (i.e. have some property in common, say Y) are included; and it is *impossible to visually imagine* that the enclosure will be included within itself (i.e. that the class can be a member of itself). One can say it in words, of course; but one cannot draw it on paper.



A class cannot be a member of itself

Thus, the idea of subsumption under a species or genus (of “Y” under Y) and that of class-membership (of “Y” in “Y”) are only partly equivalent, and should not be confused. For example, whereas the class “classes” is indeed a class, it cannot logically be a member of the class “classes.” The characterizations “is a class” and “is a member of the class ‘classes’” are not interchangeable. The class “classes” can, however, still be a member of other classes, notably of the class “not self-member classes.” Indeed, no class can be a member of “self-member classes,” because there is no such class (as just explained); and consequently, *all* classes are members of “not self-member classes.”

Note this well: even though some classes (such as the class “classes”) do *subsume* themselves, they cannot reasonably be said to be *members of* themselves. This signifies that, whatever the value of Y: ***whether the class “Y” is or is not a Y, “Y” is not a member of the class “Y,”*** i.e. it is not a member of itself. So, the proposition through which we initially defined class-membership, viz. if X is a Y, then X is a member of the class “Y,” has *a formal exception*, viz. the case where the class “Y” is intended by the term X (which is otherwise general).

(Although we have not here dealt with subclasses and overclasses, note that the said exception does not prevent the class “Y” from being a member of the *overclass* “Y-classes;” which, by definition, includes “Y” together with all the subclasses of “Y.” This is true because “Y” and “Y-classes” are at different levels, so this does not constitute self-membership. More on this below.)

Another oddity of class logic to note is that, just as “classes” is a class, “non-classes” can be said to be a class, without self-contradiction. In the case of predicate logic, such a twist is not possible; i.e. a non-class (i.e. a thing that is not a class) cannot be said to be a class: only a class can be said to be a class. This again proves the divergence between these two ways of thinking. On the other hand, note, “classes” cannot consistently be said to be a non-class; and likewise, “non-classes” cannot be said to be a non-class.

As regards the Russell paradox, it is alleged that the class of “all classes not member of themselves” constitutes a double paradox, because: if we say it is not a member of itself, then it *is a member of itself*, and if we say that it *is a member of itself* then it is not a member of itself. But it is clear, in the light of what we have realized above, that the problem lies in the consequent of the first if-then statement and in the antecedent of the second if-then statement, i.e.

in the *false* claim that they contain that “all classes not member of themselves” can be a member of itself. In truth, *no* class is a member of itself, so both the if-then statement made are wrong (non-sequiturs).

In other words, the resolution of Russell’s paradox is that the class “all classes not member of themselves,” *just like any and every other class*, is *not* a member of itself, and this claim does *not* give rise to any self-contradiction.

## 2. Why Russell’s resolutions are inadequate

Now, Bertrand Russell did formulate a somewhat similar explanation of his paradox, saying:

“No collection (whole or totality) can contain members that are defined in terms of itself; specifically, no existing collection can ever be a constitutive part of itself”<sup>42</sup>.

How does this formulation differ from my denial of the logical possibility of self-membership? Russell’s formulation obviously refers indirectly to the collection itself; i.e. he means, more directly put: a collection cannot contain itself as a member because such membership could only be triggered by *defining the collection in terms of itself*, and that is not feasible. Where is the “definition,” here? It is the term’s formulation itself, viz. “all classes that are members of themselves” in our case.

The proposed definition involves the term that is being defined; i.e. there is an element of self-reference, of

---

<sup>42</sup> Quotation from *Principia Mathematica* by Nicholas Rescher in *Paradoxes: Their Roots, Range, and Resolution*, p. 172. This “Vicious Circle Principle” was earlier formulated by Henri Poincaré.

circularity. This can be called self-definition; and it is obviously useless, if we do not already know what is intended by other means. But is definition of the class in terms of itself the issue at hand, here? I would say not: this term is merely the end-product of a reasoning process; it does not suddenly come out of the blue. It is this reasoning process that needs to be examined and evaluated.

Is the class of “all classes that are members of themselves” *produced by* self-definition? No – the underlying predicative term “all classes that *are* classes” is not problematic: indeed, all classes are classes. This predication engenders the classificatory term “all classes that *are members of* ‘classes’,” which is not per se problematic, either. What is problematic is that the latter term, using for example the fact that “classes” is a class, and so apparently a member of itself, inductively gives rise to the term “classes that are members of *themselves*.” So, the genesis of the latter, problematic term is not through mere self-definition, as Russell seems to think, but through the illegitimate traduction (putative translation) of a perfectly legitimate predication into a class-membership proposition.

The said traduction is known and proved to be illegitimate by the fact that it gives rise to the Russell paradox. My proposed resolution is easily proved with reference to the negative term “all classes that are *not* members of themselves.” If the issue were self-definition, as Russell suggests, then this term *too* would be problematic, since it also refers to itself (if only negatively); yet this term is not problematic!<sup>43</sup> Clearly, then, the issue at hand is not self-

---

<sup>43</sup> Alternatively, if the latter is thought to be problematic, then the former cannot be thought so. In any case, one of the two has to be unproblematic, for discourse to be at all possible. I mention this alternative, only because *at first sight* it may seem that the negative term “all classes not members of themselves” is the problematic one, since it apparently results in paradox (whereas the positive one

definition, but something else entirely. It is the naïve inference just described.

More accurate is Russell's concluding statement, viz. that "no existing collection can ever be a constitutive part of itself" – but that is not an explanation in itself: it is an evident consequent of the Russell paradox, i.e. a statement that there is obviously some sort of problem with any "existing collection" being "a constitutive part of itself," which implies that such a collection cannot exist. That is not an explanation, but an observation needing an explanation, and the explanation Russell offers revolves around self-definition, as we have seen. But self-definition is not, in fact, genetically involved.

I would say that, although Russell glimpsed that there is a problem inherent in the idea of self-membership (or, alternatively, perhaps, that of non-self-membership) by a class, he did not correctly understand precisely why the problem arose. He wrongly identified the problem as self-definition, i.e. as self-reference in definition. But though this feature is present and significant, it is not the crux of the problem.

In truth, the problem can only be solved through the finer conceptual analysis I have formulated. The problem arises through the assumption that subsumption under a concept and class-membership are fully parallel (effectively interchangeable) ideas – they simply are not. It is this prior faulty assumption which eventually gives rise to the Russell paradox. Once this is understood, it is easy to see why "self-membership" is unconscionable and we must say categorically that no classes are members of themselves, and in that event no Russell paradox arises anymore.

---

apparently does not). In truth, this negative term is unproblematic, and it is the positive term "all classes members of themselves" which causes the paradox. Note this well.



Later on, Russell, seemingly dissatisfied with his earlier proposed resolution of his double paradox, proposed his Theory of Types as a means to this end. In its simplest form, this theory focuses on the differences between the individual items subsumed by classes, the classes themselves, the classes of classes that the latter are subsumed by, and so on. These are successive “types,” forming a hierarchy of items from the concrete objects to more and more abstract derivatives of them.<sup>44</sup>

Very briefly put, the theory postulates that a proposition may be universally true of a lower type, but not include a higher type. For example, all classes can be said to be members of the class “classes,” without this implying that “classes” is also a member, simply because (the theory claims) it is of a higher type than all the classes it includes. So, “all classes” does not literally mean *all* classes; it can (due to “type” difference) exclude the class “classes” itself. Of course, this claim is quite arbitrary, and merely conceals the Russell paradox behind a smoke screen.<sup>45</sup>

There is some truth to this idea, but it also involves some confusion. There indeed exists a hierarchy between concrete individual items, abstract groupings of items (classes), and more abstract groups of groups of items (classes of classes). However, the Russell paradox does not arise when we try to include a class in a class of classes, i.e. a subclass in an overclass. And of course, the Russell paradox does not arise when we try to include an individual item in a class. The Russell paradox only arises when we try to include a class in itself; or (by extension) a class of classes in itself.

---

<sup>44</sup> I will not here get into a detailed analysis of this theory. I refer readers to my previous, more detailed analyses in my books *Future Logic* (chapters 43–45) and *A Fortiori Logic* (Appendix 7.5).

<sup>45</sup> I refer readers to the excellent exposition and critique of Russell’s paradox and his attempted resolutions of it, in William and Martha Kneales’ *The Development of Logic*, chapter 11.

The Russell paradox *only* arises in attempts to self-include a class, or a class of classes, *in itself*. So, hierarchy is not the issue, here, and the theory of types cannot provide a solution! It is only incidentally true that a class can only group concrete objects, or a class of classes can only group less abstract classes. The reason for the paradox is not type differences, but the fact that a class or class of classes cannot, even in imagination, include itself in itself.

Note in passing that, in our above example, the class “classes” (i.e. the class of all classes) is really a class of classes, since the units it groups are all classes. Nevertheless, it remains a class in its own right, and as such apparently belongs under itself, i.e. subsumes itself. The solution is not that it cannot subsume itself, but only that it cannot be a member of itself. This is not a matter of hierarchy (type), but a matter of changing predication into classification.

### 3. Why Rescher’s resolution is inadequate

Nicholas Rescher’s proposed resolution of the Russell paradox is somewhat but not much different from Russell’s first attempt. He applies a general principle that he has formulated, called the Successful Introduction Principle, to the case at hand. He points out that “only when something is properly identified, can it serve as a subject of meaningful discussion;” adding: “With an inappropriately identified pseudo-object... the door to contradiction and paradox is thrown wide open”<sup>46</sup>. In his view, then, the

---

<sup>46</sup> Op. cit. pp. 172-73. While this SIP principle is intuitively sound, and relevant to the case at hand, it must be said that it is not a general truth. In truth, much of human discourse is made through more or less vague terms. Relatively few terms, if any, are fully and finally defined. This is because our knowledge is essentially inductive, rather than (as many logicians imagine) deductive. We use a word (when we need one) and a working definition (when we have one) as pragmatic

Russell paradox is caused by a failure of proper identification inherent in a class defined in terms of itself.

Thus, Rescher focuses on the problem of self-definition, just as Russell does. His approach differs, as he points out, in that whereas Russell sees self-definition as an ontological issue, Rescher sees it as a communicative issue. For Russell a class cannot exist which is defined with reference to itself; whereas for Rescher, we cannot even begin to use or discuss such a class, because it is as if nothing has been said yet. Both these insights are, in my view, true. But they do not resolve the Russell paradox, as already explained. To repeat, the issue is not one of self-definition, but of passing over from a predicative discourse to a classificatory one. A class like “classes” *is* indeed a class, but “classes” cannot be *a member of* the class “classes.” These are two different systems of thought, which do not behave identically in all circumstances.

We cannot form concepts at will, just by putting some words together in a definition. A definition is a hypothesis subject to verification by logic. In the case of the definition of class-membership, viz. the proposition that “if X is a Y, then X is a member of the class ‘Y’,” double paradox is found to ensue in the special case of “if ‘Y’ is a Y, then ‘Y’ is a member of ‘Y’.” Therefore, the rule we tried to formulate (i.e. the said definition of class-membership) must have an exception; namely, when X stands for the class “Y.”

---

tools. Our understanding of their reference is at all times tentative and temporary. As our knowledge grows and our analyses are refined, our terms (hopefully) gain increasing accuracy, become more pointed. So, the principle here proposed by Rescher should not be taken too generally. Meaningful discussion occurs all the time with imperfectly defined words; that is the norm rather than the exception, and we usually manage very well to avoid contradictions nonetheless. Meaning is not essentially a verbal issue, but one of insight and understanding.

The Russell paradox is valuable in that it teaches us that an exception must be allowed for in the defining formula. We initially assume the general definition to be valid; but after we find out that it leads to self-contradiction and paradox, we are obliged to retract it in part. We are logically forced to altogether reject the notion of self-membership. This is a very limited and precise retraction, but it restores consistency to class logic. Our hypothetical definition has been adjusted to avoid inconsistency. This is the correct resolution of the Russell paradox.

#### 4. Other incarnations of the paradox

It should be mentioned that Russell realized that the concept of class was at issue somehow, but then he tried to show that the same paradox could be formulated within predication. As Kneale and Kneale rendered it<sup>47</sup>:

“Instead of the class which is supposed to contain all classes that are not members of themselves let us consider the property of being a property which does not exemplify itself. If this property exemplifies itself, then it cannot exemplify itself; and if it does not exemplify itself, then it must exemplify itself. Clearly, the nature of the trouble is the same here as in the original paradox, and yet there is no talk of classes.”

The implications here are that some properties “exemplify” themselves, while some do not. That is, a property Y: in some cases, Y is Y (i.e. has property Y) and in others Y is not Y (i.e. lacks property Y). If Y is Y, then Y exemplifies itself; and if Y is not Y, then Y does not exemplify itself.

---

<sup>47</sup> *The Development of Logic*, p. 655.

Is it, however, ever true that Y is Y? Can anything be pointed to that has precisely itself as a predicate? I would say not. Clearly, 'Y is Y' here does not refer to a predicative tautology, like 'white is white'. Rather what we mean by 'Y is Y' here is that 'whiteness is white', i.e. *Y-ness* (*the property Y*) is Y (has the property Y). This is a very different statement, which does not follow from the preceding one.

And the same critique of it can be brought to bear for it as for the classificatory version. We may well say that if X is Y, then X "exemplifies" Y-ness (i.e. is one of the things that has Y-ness), but in the special case where X is replaced by Y-ness (i.e. the property Y), we must make an exception, and deny that 'Y-ness is Y' implies 'Y-ness exemplifies Y-ness (i.e. itself)'. To avoid an eventual Russell paradox, we must at least short-circuit such implication, if not totally reject the consequent (saying that it is never true that Y-ness exemplifies Y-ness), if not even reject the antecedent (saying that it is never true that Y-ness is Y).

Russell claimed that if "the property of non-self-exemplification" exemplifies itself, then it lacks "the property of non-self-exemplification" and so cannot exemplify itself; and if does not exemplify itself, then it has "the property of non-self-exemplification" and so must exemplify itself. To my mind, this means we must reject the very idea of self-exemplification and say that *nothing* exemplifies itself (even if we happen to find some Y-ness which is Y). In that event, we can well say that "the property of non-self-exemplification" does not exemplify itself, without any contradiction arising.

The same thing can be said using other terms: some concepts are "instances" of themselves, while some are not. That is, we can define instantiation by saying: if X is a Y, then X is an instance of Y (i.e. X is an individual within the

group called Y); but we must make an exception in the special case where X is replaced by Y, and deny that if Y is a Y then Y is an instance of Y. It is clear that Y here refers specifically to a concept, and not just any kind of thing. An example would be “the concept ‘concept’ is a concept; therefore, the concept ‘concept’ is an instance of itself.”

This is clearly very similar to saying “the class of all classes is itself a class; therefore, the class of all classes is a member of itself;” and a similar paradox is bound to emerge from it, which can be neutralized in a similar manner. That is, we must upon reflection say that *nothing* is an instance of itself, so that in fact “the concept of non-self-instantiation” is not an instance of itself. A similar line of reasoning can be followed with regard to “inclusion” and any other similar relations.

Clearly, none of these issues relate to predication as such, but to more geometrical ideas (classifying, having/belonging, instantiating, including) through which we try to represent predication. These relations are intended as metaphors for predication; but it turns out that these analogies are not perfect. The lesson the Russell paradox teaches us is that these derivative relations all have exceptions (as already detailed above). The Russell paradox does not affect predication as such; predication as such is immune to it. Although Russell was right to investigate whether his paradox meme affects predication as such, he was wrong to conclude that it does; it does not.

## 5. About the Barber paradox

There is a town with only one barber. Some men in it shave themselves; the rest do not, but are all shaved by the barber. What about the barber? Does he shave himself or not? If he does, he is one of the men who shave themselves (no problem); if he does not, he is one of those shaved by the

barber, i.e. by himself (apparent contradiction). Obviously, in his case, the two classes of men overlap: he both “shaves himself” and he is “shaved by the barber.” This is hardly problematic: it just tells us that the supposed either-or disjunction between the two groups of men is not really one; his case is a logically implied exception to it. That is, in his case, unlike in other cases, being shaved by the barber does not imply not shaving himself.

How does this compare to the Russell paradox, of which the barber paradox is touted as an illustration? I resolve the Russell paradox by saying that no class is a member of itself; it follows that we can say that the class of “classes not members of themselves” is necessarily not a member of itself, without this giving rise to a contradiction. I resolve the barber paradox differently, by saying that, unlike all other men in town, the barber can well both shave himself and be shaved by the barber (himself). In the Russell paradox, one of the conflicting classes is wholly eliminated. In the barber paradox, the conflict between the classes is softened in a single case (that of the barber himself). Since the two paradoxes are not resolved in the same way, they must be regarded as logically distinct.

## 10. Hempel's paradox of confirmation

*From Logical and Spiritual Reflections 1.8.*

### 1. Traditional analysis

Carl Gustav Hempel<sup>1</sup> in the 1940s exposed an alleged “paradox of confirmation”, which suggested that a fully consistent formal inductive logic is impossible. This is commonly called “the raven paradox”, and may be described as follows:

a) The observation that Some ravens are black (Some A are B) confirms the hypothesis that All ravens are black (All A are B).

The latter proposition may be contraposed to All non-black things are non-ravens (All nonB are nonA).

b) Next, consider the observation that Some apples are green (Some C are D). This is convertible to Some green things are apples (Some D are C).

It follows from this proposition that Some non-black things are non-ravens (Some nonB are nonA), since green things are not black and apples are not ravens.

Now, just as Some ravens are black (Some A are B) confirms the hypothesis that All ravens are black (All A are B), so Some non-black things are non-ravens (Some nonB are nonA) confirms the hypothesis that All non-black things are non-ravens (All nonB are nonA).

---

<sup>1</sup> Germany-USA, 1905-97.



This induced proposition may in turn be contraposed to All ravens are black (All A are B), and here lies the difficulty, for it appears that the mere observation of some green apples is enough to confirm the hypothesis that All ravens are black! Note well that to achieve this result we did not even need to observe any black ravens.

c) It follows from the preceding that we can equally well, using the same logical process, given Some apples are green, confirm the hypothesis that All ravens are pink, or any other color (except green) for that matter.

d) This is in itself a mystery: how can apples tell us about ravens? Intuitively, this has to be viewed as a *non-sequitur*. Moreover, in the case of black ravens, the existence of black ravens has empirical backing, as already indicated; so, the 'inference' from green apples to All ravens are black still seems somewhat reasonable. But in the case of pink ravens, we have never observed any such animals; so, the 'inference' from green apples to All ravens are pink seems quite unjustifiable.

Moreover, knowing by observation that Some ravens are black, how can we 'conclude' that All ravens are pink? Even if we do not claim all ravens black, but only claim all ravens pink, we would in such circumstances be upholding contrary propositions, namely the particular one that some ravens are black and the general one that all ravens are pink.

Moreover, even if we have never observed the color of any ravens, we can according to the above inductive process simultaneously conclude many contrary statements such as All ravens are black, All ravens are pink, All ravens are orange, etc. This too is a result that flies in the face of the law of non-contradiction.

Furthermore, the same can be done with reference not only to green apples, but also to apples of other colors (except black or pink, etc. as the case may be), and indeed to things (non-ravens) other than apples. In that event, almost

anything goes in knowledge, and *we can at will affirm or deny just about anything about just about everything!*

This then, according to traditional presentations<sup>2</sup>, is Hempel's paradox. It appears, from such analysis that the inductive processes of confirming hypotheses (such as generalizations directly from experience or indirectly from logical derivatives of experience) are fundamentally flawed. The analysis involved is quite *formal*, i.e. it can be performed in terms of symbols like A, B, C, D – and so it has universal force.

It follows that induction is bound to result in various absurdities: apparent non-sequiturs, many contradictions, and ultimately imply the arbitrariness of all human knowledge. Clearly, Hempel discovered here a serious challenge to inductive logic and logic in general.

## 2. Novel analysis

As I will now show in detail, *the above analysis is inaccurate in some important respects*. I will show that although Hempel did indeed discover an interesting formal problem for logicians to consider and solve, this problem does not result in what we would call a paradox. That is, there are valuable lessons to be learned from Hempel's paradox (as we may continue to call it conventionally), but it does not present logic with any insurmountable predicament.

a) The first operation described above is the commonly used inductive process of *generalization*. A particular proposition (Some A are B) is turned into a general one (All A are B). The particular supports the general in the way that positive evidence confirms a hypothesis. Their logical relation is adductive. 'Some' here means 'at least

---

<sup>2</sup> See for instance: the article in Wikipedia at [http://en.wikipedia.org/wiki/Raven\\_paradox](http://en.wikipedia.org/wiki/Raven_paradox).

some, possibly all and possibly only some' – and by generalizing we are opting for the hypothesis 'all' in preference to the hypothesis 'only some'.

However, it would be an error to consider that Some A are B is *alone* capable of inductively justifying All A are B. Such generalization is an inductively permissible inference *provided we have looked for and so far not found any A that are not B*. For if we *had* found (by direct observation or by some reasoning) that Some A are not B, we would certainly not have generalized. Moreover, if we *later* do come across an A that is not B, we would have to *particularize* All A are B back to Some A are B.

This condition *sine qua non* of generalization, viz. to remain on the lookout for contradictory instances and adjust one's judgment accordingly, is not stressed or even mentioned in the earlier presentation, note well. Yet this is a known and accepted rule of scientific thought at least since the time of Francis Bacon, who emphasized the importance of the "negative instance" in induction. To ignore this condition is bound to lead to contradictions sooner rather than later.

Regarding the contraposition of All A are B to All nonB are nonA, it is of course a deductive act. Even so, we must keep in mind that the conclusion All nonB are nonA is only due to the prior inductive inference of All A are B from Some A are B. No observation is required for the deduction, but we remain bound by the need to keep checking the previous inductive act, i.e. to remain alert for eventual cases of A that are not B.

b) Now, let us grant that Some C are D, as above. Some C are D readily converts to Some D are C. However, Some D are C does not formally imply that Some nonB are nonA – some syllogistic inference is tacitly involved here, which ought to be brought out in the open. Clearly, we tacitly take for granted the premises that *green is not black* and *apples*

*are not ravens*, whence: the following two successive syllogisms are constructed:

1st figure, EIO:

No green thing is a black thing (No D are B)

Some apples are green (Some C are D)

Therefore, Some apples are not black (Some C are not B).

3rd figure, AII:

All apples are non-ravens (All C are nonA)

Some apples are non-black (Some C are nonB)

Therefore, Some non-black things are non-ravens (Some nonB are nonA).

Whence, by generalization we obtain: All non-black things are non-ravens (All nonB are nonA); and then by contraposition: All ravens are black (All A are B). Note that the premises that led to this general conclusion do not include Some ravens are black; i.e. this conclusion is based on no empirical observation of black ravens.

Note too that we could have obtained the same result with the premises No ravens are green (No A are D) and No apples are black (No C are B). Note also that, though the syllogisms involved are deductive processes, all such tacit premises require prior observations and generalizations (i.e. inductions) to be adopted.

Moreover, it is significant to note that these syllogisms could not be constructed if the colors of the ravens and apples under consideration were the same (both green or both black), or if ravens and apples were not mutually exclusive classes. We also assume here that a raven cannot have more than one color (e.g. be partly black and partly green or whatever, or sometimes the one and sometimes the other); and similarly for an apple.

The next step was to generalize Some nonB are nonA to All nonB are nonA. But here again, generalization is allowed only provided we have no evidence or inference from any other source that Some nonB are not nonA. That is, in our example, we must remain conscious that it is

possible that some non-black things are not non-ravens, i.e. are ravens, which means we might yet find some non-black (albino) ravens out there.

Here too, we must make sure, in accordance with Bacon's crucial principle of adduction, that there is no conflicting observation that obstructs our expansive *élan*. This is all the more necessary, since here the premise of generalization Some nonB are nonA was obtained indirectly by deduction from previous products of induction, whereas previously our premise Some A are B was (supposedly) directly observed.

Note further that these two generalizations have the regulatory conditions that Some A are not B or Some nonB are not nonA, respectively, not be found true – and these conditions are one and the same since these two propositions are logically equivalent by contraposition. This means that in either case, whether we reason directly from black ravens or indirectly from green apples, there is the same implicit condition for generalization – that in our experience or reasoning to date no non-black ravens have appeared.

Thus, whichever of these two generalizations we opt for, the condition that there be no known instances of A which are not B is unaffected, and the dependence of the truth of All A are B on this condition is unchanged. Note too, the same condition holds before and after such generalizations. That is, even after such inductive process, if we discover new evidence to the contrary, we logically may and indeed must *retract*.

As previously stated in c) and d): using the same logical process, given Some apples are green, we can equally confirm the hypothesis that All ravens are pink, and many other wild hypotheses that conflict with each other<sup>3</sup>.

---

<sup>3</sup> To show propositions with different predicates are in conflict, we use syllogism. For instance, All ravens are black and All ravens are

Obviously, we are doing something wrong somewhere, and have to take action to either prevent such absurd eventual consequences or correct them when and if they occur. I will now explain the solution to the problem.

Generalization is never an irreversible process. So, if any generalization leads to contradictions, we are free and indeed obligated to particularize. The question of course remains: *in what precise direction and how far back should we go?* Still, what this means is that there is no ‘paradox’ in inductive logic as there is in deductive logic; almost everything (with the exception of logic itself – especially the laws of thought on which it is built) is and ever remains ‘negotiable’.

In deduction, a contradiction is a far more serious event, because the process leading up to it is presumably necessary. But *in induction, we know from the outset that the connection between premise(s) and conclusion is conditional – so contradictions are expected to arise and it is precisely the job of inductive logic to determine how to respond to them.*

Dealing with contradictions is a branch of inductive logic, called **harmonization** or conflict resolution. This is not something rare and exceptional – but occurs all the time in the development of knowledge. *Sometimes conflicts are resolved before they take shape, sometimes after. If we see them coming, we preempt them; otherwise, we perform the possible and necessary retractions.*

Particularization of a general proposition is retraction. More broadly, retraction means rejection or modification of a theory in the light of new evidence. Thus, for example: till now, I have seen only black ravens, and assumed all are black; tomorrow, I may notice a white raven, and change my view about the possible colors of ravens.

---

pink are incompatible, because knowing that No black things are pink, we obtain, by syllogism (1<sup>st</sup> figure, EAE): No ravens are pink, which is contrary to All ravens are pink.

Hempel is evidently or apparently unaware of this crucial aspect of inductive reasoning, else he would not have viewed contradictions arising in the course of induction as paradoxical. Nevertheless, the situation described by him is interesting in this context, for reasons he did not (I think) realize.

For after the first generalization, starting from Some ravens are black (Some A are B), if we belatedly discover that Some ravens are not black (Some A are not B), *we simply return to our initial observation* that Some ravens are black (Some A are B). Whereas after the second generalization, starting from Some non-black things are non-ravens (Some nonB are nonA), if we belatedly discover the same conflicting evidence, we cannot simply deny All ravens are black (All A are B).

Why? Because *this would still leave us with part of our generalization*, viz. the claim that Some ravens are black (Some A are B). That is to say, we would expect 'All A are B' plus 'Some A are not B' to yield the harmonizing conclusion 'Some A are B and some A are not B'. The negative particular does not eliminate the positive particular underlying the positive generality; since we previously (due to said generalization) believed the generality, we now have a leftover to account for.

In the case of All ravens are black, such retraction is not noteworthy, since we know from experience Some ravens are black; but in the case of All ravens are pink, we have a serious problem, for there is no shred of evidence for a claim that Some ravens are pink! In other words, the proposed retraction cannot suffice in the situation presented by Hempel, i.e. when All A are B is induced from Some nonB are nonA.

### 3. The problem and its solution

In my view, this is the crux of the problem revealed by Hempel's exploration. The problem is not exactly a paradox, since *the validity of generalization formally depends on such process not giving rise to any eventual contradiction*.<sup>4</sup>

That from the observation of some green apples we may by generalization infer that All ravens are black and All ravens are pink and many other conflicting conclusions – this is amusing, but not frightening. For in such situation of self-contradiction, we can by retraction find ways to harmonize our knowledge again. The problem is temporary.

On the other hand, what Hempel has here uncovered is that we cannot always retract simply by particularization of the conflicting theses. Particularization seems acceptable in some cases (e.g. with black ravens), but in other cases it yields unacceptable results (e.g. with pink ravens), because the logical remainder of such retraction is devoid of empirical basis.

Suppose, using Hempel's method, starting from green apples, we induce both the generalities All ravens are pink and All ravens are orange. These two conclusions are in conflict. Let us say we decide to resolve the conflict by denying them both; that still leaves us with two

---

<sup>4</sup> A paradox is a thesis that formally contradicts itself or deductively leads to contradictory propositions. From a single such paradox, we may conclude that the thesis in question is false; logic as such is not put in question, because the contradiction involved is merely conditional. A double paradox, on the other hand, is a serious threat to logic; here, both a thesis *and* its contradictory are paradoxical, so the contradiction is *unconditional*. In that case, logic cannot declare either of them true or false – but must among them find either a non-sequitur (as in the Barber paradox) or a meaningless term (as in the Liar paradox). That is, logic must challenge either one or more of the implications involved, and/or one or more of the terms or theses involved. The Hempel scenario does not give rise to an unconditional/double paradox.



propositions Some ravens are pink and Some ravens are orange.

These two particular propositions are not in conflict – and, let us take for granted, neither of them has any empirical basis, yet they both got somehow cozily ‘established’ in our knowledge! They were introduced by the generalizations from green apples, yet they were not dislodged when we abandoned the corresponding generalities. We are stuck with them, even though the complex processes that led to them have been revoked.

It is unthinkable that such *particulars* (whether true or untrue) should emerge from the unrelated observation of green apples (or whatever else). This I believe is the significant problem uncovered by Hempel. The problem is not the conflict of generalities or between general and particular propositions, so it is not about paradox. The problem has to do with ‘collateral damage’ to knowledge, through incomplete correction of errors.

I suggest the following solution for it: when we generalize from Some A are B to All A are B, and then discover that Some A are not B, we particularize All A are B back to Some A are B. That is normal procedure, which we all commonly practice.<sup>5</sup>

On the other hand, when we obtain All A are B by generalization from Some nonB are nonA to All nonB are nonA (followed by contraposition of the latter), then when we discover that Some A are not B, we cannot merely particularize All A are B back to Some A are B, but must also retract the intermediate premise of the proposition All A are B, viz. All nonB are nonA, and return to Some nonB are nonA.

In view of the latter retraction, we in fact no longer have a basis for claiming Some A are B (this cannot be deduced from Some nonB are nonA). It would be an error of

---

<sup>5</sup> Symbolically,  $A + O = IO$ .

induction to forget the actual *source* of our belief in All A are B. The distinction between the inductive grounds Some A are B and Some nonB are nonA must be kept in mind, so that in the event of discovery of contradictory evidence, viz. that Some A are not B, we particularize back to our *exact same previous position* in each case.

We may thence formulate the following new law of inductive logic, which may be called *the law of commensurate retraction*: a product of generalization like All A are B cannot be treated without regard to its particular source; when if ever it is denied by new evidence, we must retreat to the same initial particular and not to *some other* particular that was implied by the generality when it seemed true but is now no longer implied by it since it is no longer true.

In other words, when and if we come upon a contradiction of the sort considered here, we must realize that this does not merely discredit the generality that was previously induced, but more deeply discredits the inductive act that gave rise to it. Thus, we should not retract by mere particularization, but carefully verify whether the remaining particular has any independent basis and if it has not we should return far back enough to the *status quo ante* to make sure no unconfirmed particular remains.

This seems like a perfectly reasonable instruction – to reverse and clean up all traces of an inductive act that was found illicit, i.e. that led us into a logical impasse.

All this means that, using ordinary procedures of logic, we would never fall into a self-contradictory situation (e.g. claiming paradoxically All ravens are black and All ravens are pink). The fact that generalizations may yield incompatible results is commonplace; we daily deal with such conflicts without difficulty. When such conflicts arise, we are logically required to harmonize. If we cannot find a specific way to resolve the conflict, the conflict is

resolved in a generic manner, viz. all the generalizations involved are put in doubt.

In a situation where two or more propositions are put in doubt by mutual conflict, we would naturally give more credence to one that has some direct empirical basis (like All ravens are black) than to one that merely emerged from indirect projection (like All ravens are pink). We need not treat all conflicting propositions with equal doubt, but may be selective with regard to their inductive genesis.

With regard to the evidence for conflicting thesis – obviously, if we have no data on black or pink ravens, we would not know which way to retract, and both generalizations would be problematic. But if we have observed some black ravens and never observed any pink ones, we would naturally opt for the generalization that All ravens are black (All A are B). On the other hand, if we have observed both black ravens and pink ravens, we would make neither generalization and simply conjoin the two particulars.

With regard to the inductive processes used – direct generalization would naturally be favored over the indirect sort envisaged by Hempel. If the conflict at hand can be resolved by ordinary means, e.g. with reference to empirical considerations, we need not bother to backtrack with reference to process. But in cases where we have no other means of decision, process would naturally be the focus of revision.

A possible objection to the law of commensurate retraction would be that in practice we rarely manage to keep track of the exact sources of our generalizations. Such ignorance could conceivably occur and cause some havoc of the type Hempel described in our knowledge.

However, we may also point out that *in practice we just about never find ourselves in the situation described by Hempel*. How often does anyone generalize from a proposition like Some nonB are nonA? The statistical

answer is ‘probably never’ – Hempel’s paradox is just *a remote formal possibility* that logicians have to consider, but its practical impact is just about nil.

Moreover, we are not likely to arrive at a proposition of the form Some nonB are nonA, except by the sort of reasoning above depicted, i.e. through some other terms like C and D. We cannot directly *observe* that Some nonB are nonA. Observation relates primarily to positive phenomena; it can be about negative phenomena but only indirectly. This suggests that if we did encounter a situation of Hempel paradox, we would likely be aware of how it arose.

Another remark worth making is that the above solution of the problem raised in Hempel’s paradox can be characterized as heuristic; it is repair work by trial and error. But I have already proposed in my work *Future Logic*<sup>6</sup> a detailed, systematic, formal treatment of induction, by means of factorization and formula revision. I believe that is free of the Hempel’s problem, since every formal possibility is included in the factorial formulas developed.

With regard to solutions to Hempel’s paradox offered by other logicians, e.g. those by Goodman and by Quine described in the earlier mentioned Wikipedia article:

“Nelson Goodman suggested adding restrictions to our reasoning, such as never considering an instance as support for ‘All P are Q’ if it would also support ‘No P are Q’ ... Goodman, and later another philosopher, [W.V.] Quine, used the term *projectible predicate* to describe those expressions, such as *raven* and *black*, which *do* allow inductive generalization; *non-projectible* predicates are by contrast those such as *non-black* and *non-raven* which apparently do not. Quine suggests that it is an empirical question which, if any, predicates are projectible; and notes

---

<sup>6</sup> First published in 1990, a few years before Hempel’s death. See part VI.

that in an infinite domain of objects the complement of a projectible predicate ought always be non-projectible. This would have the consequence that, although "All ravens are black" and "All non-black things are non-ravens" must be equally supported, they both derive all their support from black ravens and not from non-black non-ravens."

I find these proposals reasonable and not incompatible with my own. However, I think mine is a little more precise in pinpointing the problem at hand and its solution.

Goodman's suggestion to restrict induction from a proposition if such process yields conflicting conclusions is logically sound. Only his instruction cannot be obeyed preemptively, but only *after* we discover that the process yields conflicting conclusions. So, it is not a preventative, as he seems to consider it, but an after the fact correction. It can therefore be regarded as about the same as the law of commensurate retraction I above propose. The only difference is that he does not seem to have made a distinction between the conflict of generalities and the underlying leftover particulars.

As for "non-projectible predicates", I would agree that negative terms (complements) present a general problem in induction. Although deductive logic makes no distinction between positive and negative terms, phenomenology does distinguish between the presence of positive phenomena and their absence. Whereas we can observe a positive phenomenon (like a black raven) without regard to its negation, we cannot mention a negative term (like non-black or non-raven) before thinking of and looking for the corresponding positive phenomenon and failing to find it.<sup>7</sup> Thus, a truly negative term can never be truly empirical. Its content is never 'I have seen something', but always 'I have diligently looked for something and not found it'. A negative is 'empirical' in a lesser, more derivative sense

---

<sup>7</sup> See my essay on this topic in *Ruminations* (part I, chapter 9).

than a positive. It already involves a generalization of sorts, from ‘could not be found’ to ‘was not there to be found’. It follows from this insight that generalization from negative terms, such as Some nonB are nonA, can only proceed with unusual caution and skepticism. Hempel’s scenario further justifies such tentativeness. We are might even be tempted as a radical solution to simply always interdict generalization for a truly negative subject. If any manner of discourse has certain likely illogical consequence, logicians are wise to formulate a preemptive law of logic of this sort.

Another temptation is to deny any meaningful content to propositions of the form Some nonB are nonA. Such a proposition is formally implied by All A are B, and compatible with Some A are B, No A are B and Some A are not B – but does it really tell us anything? Indeed, since nothing can be inferred about A or B (as subjects) from Some nonB are nonA, what information does such a proposition contain? Could one not conceivably assert such a proposition using any almost two terms taken at random? This sort of doubt could be used to further justify interdiction of generalization from such propositions.<sup>8</sup>

However, since a less radical solution, namely the above-proposed law of commensurate retraction is possible, we perhaps need not go so far. Rather than preemptively forbid certain doubtful processes under all conditions, I prefer to allow them in case they occasionally work, and prepare the

---

<sup>8</sup> These questions are made clearer if we consider the eventual negation of Some nonB are nonA, i.e. the form No nonB is nonA, which implies All nonB are A. In the event the latter proposition is true, we would have a negative term (nonB) included in a positive (A). This could be taken to mean that almost all the world (except things that are B) falls under A. For this to happen, A would have to be a very large concept. Such a concept would be very exceptional and almost meaningless. Whence, we can say that Some nonB are nonA is almost always true, and at the same time not very informative.

appropriate corrective mechanism for when they fail to work.

To sum up, I believe we have convincingly shown here that Hempel's so-called paradox does not present the science of logic with any insuperable difficulty; it is made out to be a bit more daunting than it really is. Even so, it is an interesting contribution for logicians to ponder over.

## 11. Goodman's paradox of prediction

*From Logical and Spiritual Reflections 1.9.*

### 1. The alleged problem

Nelson Goodman<sup>56</sup> proposed in 1955 a “riddle of induction” (as he called it<sup>57</sup>) or “paradox of prediction” (as others have characterized it), which seemed to demonstrate a formal difficulty in generalization. This may be stated as follows:

“Goodman ... introduce[d] the color grue, which applies to all things examined before a certain time  $t$  just in case they are green, but also to other things just in case they are blue and not examined before time  $t$ . If we examine emeralds before time  $t$  and find that emerald  $a$  is green, emerald  $b$  is green, and so forth, each will confirm the hypothesis that all emeralds are green. However, emeralds  $a, b, c, \dots$  etc. also confirm the hypothesis that all emeralds are grue. In this case emeralds  $a, b, c, \dots$  examined after time  $t$  should be grue, **and therefore blue!**” (Emphasis mine)<sup>58</sup>

---

<sup>56</sup> USA, 1906-98.

<sup>57</sup> Or more pretentiously, “the new problem of induction”.

<sup>58</sup> Here I'm quoting: [http://en.wikipedia.org/wiki/Nelson\\_Goodman](http://en.wikipedia.org/wiki/Nelson_Goodman). Elsewhere, we are informed that “applies to all things examined before  $t$  just in case they are green but to other things just in case they are blue” is Goodman's own wording in his original presentation in *Fact, Fiction, and Forecast* ([http://en.wikipedia.org/wiki/Grue\\_%28color%29](http://en.wikipedia.org/wiki/Grue_%28color%29)).



The significance of this artifice, according to its proponents, is that although green and "grue" have the same linguistic form, and so should be subject to the same logical processes (in this case, the inductive process of generalization), they are internally quite different types of concepts, since the first implies a similarity between its past and future instances, while the second suggests a change of color over time, so that the result is paradoxically quite different if we generalize with reference to the one or the other.

## 2. The logical solution

However, as I shall now formally demonstrate, this is merely a sleight of hand, for though the act of generalization is equally valid for green and for grue, it does not follow that we can infer any emeralds to be blue from the induced general proposition that all emeralds are grue. That is to say, the conclusion "and therefore blue" in the above presentation is an erroneous deduction.

To expose this simple error, the given scenario must be reformulated more carefully (the symbols X, A, B, C are mine):

- Say we examine all available emeralds (X), till a certain time (t), and finding them all to be green (A), we ordinarily conclude by generalization that All emeralds are green (All X are A), although we know [from past experience with induction in general] that the next emerald we find, after time t, might well turn out to be blue (B) [or indeed, to be some other color<sup>59</sup>]!

---

<sup>59</sup> This is my own interpolation, to make Goodman's thesis more accurate. For there is no reason to suppose a priori that only blue emeralds might eventually be found. We are only guessing the possibility of blue emeralds, not basing it on any specific observations – therefore any other color is equally probable (or improbable).

- Let us now following Goodman introduce a new concept “grue” (C) to be defined as the class grouping all things that were examined before a certain time *t* and found to be green (A) and all things *not* examined before time *t* which happen to *be* blue (B) [or indeed, to be some other color].<sup>60</sup>
- Applying this definition, all X (emeralds) examined before *t* were found A (green) are also C (grue); i.e. by syllogism we can infer Some X are C. As for remaining eventual cases of X, those not examined till after time *t* [if ever], *each will either be found to be A (green) or to be B (blue) [or indeed, to be some other color]*; in that sense, the latter X too are C. Hence, All X are C would seem a reasonable conclusion.
- But it certainly does not logically follow from the preceding that any emeralds will indeed be found to be any color other than green, i.e. that any X are B [i.e. blue, or whatever non-green color]! For, properly understood, the category C is *not* formulated as a disjunction of A or B that is bound to actualize both some cases of X-A and some cases of X-B.
- If you look closely, you will see that C includes on the one hand things already *known* to be A (green emeralds already observed) and on the other hand a palette of things of still *unknown* qualification, i.e. either A or B (blue) [or even some other color]. The latter is a disjunction of

---

Nevertheless, my refutation of Goodman works just as well without this added comment.

<sup>60</sup> Note that the latter things are stated to be merely “not examined until time *t* [yet, if ever]”; this is not to be confused (as some commentators have done) with “examined after time *t*”, for no matter how many things we do eventually examine, we will obviously never achieve (or know we have achieved) a complete enumeration of all such things in the universe. Note also that the concept grue is here defined as a general predicate for any eventual subject (“things”), rather than specifically for emeralds.

*conceivable* outcomes, not one of *inevitable* outcomes. To infer X-B as an *actual* outcome would therefore be a *non-sequitur*.

- The fact that we do not know whether any future X will be found A does not allow us to infer from this disjunction of possibilities that some future X will necessarily be B. We do not yet know whether any future X will be found B, either. We may well find that All X are A (All emeralds are green) remains forever applicable after time t as before time t (as predicted in the initial ordinary generalization).
- The premises 'All X are C' and 'All C are A or B' indeed yield the syllogistic conclusion 'All X are A or B'. But the disjunction 'A or B' here cannot be interpreted differently in the major premise and in the conclusion. The disjunction in the premise not being extensional, the disjunction in the conclusion cannot be treated as extensional<sup>61</sup>. To do so would be to commit the fallacy of four terms.

It is thus clear from our exposition that *the introduction of the concept "grue" has changed nothing whatsoever in the inductive possibilities offered by the given data*. The correct inductive conclusion remains unaffected by Goodman's fun and games. All Goodman has succeeded in doing is artfully conceal his fallacious *deductive* reasoning (misinterpretation of the kind of disjunction involved); it is all just sophistry.

In the thick smoke of Goodman's rhetoric, it is made to appear as if blue emeralds are as easy to predict as green ones. But that is not at all the logical conclusion according

---

<sup>61</sup> That is, a base of the given disjunction is Some C *might be* B, whereas the corresponding base of the allegedly inferred disjunction is Some X *are* B. But to imagine something happening is not proof it has to in fact happen sometimes. The conclusion does not follow from the premise.

to inductive logic. Why? Because in the case of the hypothesis that future emeralds observed will be found green, we have some concrete data to support it, namely that all present and past emeralds observed have been found green.

Whereas, in the support of Goodman's hypothesis that blue emeralds will appear, we have no experiential evidence whatever so far. All we can say is that it is not inconceivable that blue emeralds might one day be found, but that does not imply that any ever will. 'Not inconceivable' does not justify actual prediction. It just means 'imaginable in the present context of knowledge'.

That is, all we have is a general epistemological principle to remain open-minded to all eventual outcomes, based on past experience relating to all sorts of objects, that novelty does appear occasionally. But such scientific open-mindedness is not equivalent to a positive prediction of specific changes. It is just a call, in the name of realism, to avoidance of prejudice and rigidity.

### **3. The concept of 'grue'**

A question we ought to ask is whether Goodman's "grue" construct is a well-formed concept?

An ordinary concept of "grue" (or green-blue) would simply be formulated as "green and/or blue". We may well find it valuable to introduce such a concept, perhaps to stress that green and blue are close in the range of colors, or that some things are partly green and partly blue, or sometimes green and sometimes blue, or that some hues in between are hard to classify as clearly green or clearly blue. The dividing line between these colors is after all pretty arbitrary.

Given that some emeralds are green, we could then deduce that some emeralds are grue. It would be equally valid to

induce thence that all emeralds are green or that all emeralds are grue. This would imply no inherent self-contradiction, because to say that all emeralds are grue does not imply (or exclude) that any emeralds are blue. All emeralds are grue is formally compatible with the eventuality that all emeralds are green. So, there is no "paradox of prediction" in fact.

Goodman's "grue" construct is no different from this ordinary concept with respect to such logical implication. Its difference is not in the involvement of disjunction (green or blue), since such disjunction is quite commonplace; for example, the concept "colored" means (roughly) "red, orange, yellow, green, blue, indigo or violet". The significant difference in Goodman's construct is its involvement of temporal-epistemic conditions. This serves the rhetoric purpose of clouding the issues.

Defining the concept "grue" as the class of all things examined before time  $t$  and found to be green and all things *not* examined before  $t$  that happen to *be* blue – involves a self-contradiction of sorts. If I have not yet examined the things after time  $t$ , how can I positively say of any of them that they are blue? I could only make such a statement *ex post facto*, after having examined some of the things after time  $t$  and found them blue.

Alternatively, it would have to be said *by a 'third party' looking on*, who has examined some of the things before time  $t$  and found them blue, and who is observing my situation before I have done the same. But as regards all current observers taken together, they cannot logically adopt such a hypothesis, about things that happen to be blue although they have not yet been observed to be so. We can only consistently talk about things that *might yet* be found blue. For this reason, Goodman's grue concept is not well-formed.

Grue is primarily defined as the union of green things and blue things; but it does not follow from such definition that if some things (such as emeralds) are green, then other *such* things (i.e. other emeralds) must be blue. To say that a kind of thing (emeralds) is grue is not to intend that its instances must cover the whole range of possibilities included under grue. The concept of grue remains legitimate provided we find the predicates it collects together (green, blue) scattered in various kinds of thing (emeralds, the sea, etc.). Thus, every ordinary predicate involves some uncertainty as to its application to specific subjects. Moreover, this applicability may vary with time: according to our context of knowledge, and according to changes occurring in the objects observed. Therefore, there is no need to involve such epistemic and temporal factors in the definition of any of the concepts we propose. Such factors are inherent to conceptualization.

The reason Goodman introduced such complications in his definition of “grue” was because he wanted to refute (or give the impression he was refuting) the process of generalization we commonly use to develop our knowledge on the basis of limited observation.

According to inductive logic, observing that some X are A, and so far seeking and not finding any X that are not A, we may generalize and say All X are A. This remains effectively true for us so long as we have no evidence of any X that is not A. Generalization involves prediction, i.e. saying something about cases of X we have not yet observed and maybe never will.

Goodman wished to demonstrate that we are equally justified in predicting a negative outcome (i.e. not A, e.g. B) as a positive outcome (i.e. A)<sup>62</sup>. He did not realize the

---

<sup>62</sup> To do so, he needed to construct a concept that would include both A and notA, so that generalization could be formally shown to be able to go either way. However, since a concept including

logical *justification* of our generalizations<sup>63</sup>. We are not arbitrarily predicting that the cases of X we observe in the future will be A rather than not A. We are just sticking to *the same polarity* (A), because it is the only polarity we have any empirical evidence for so far. Comparatively, to predict *the opposite polarity* (not A, in this context) would be purely arbitrary – a wild assertion. Specifically for X, the first move has some empirical support, whereas the second has none at all.

Goodman simply did not realize this difference in justification between the two courses, though it is obvious to anyone who takes the time to reflect. He thus failed to apply the inductive principle that a confirmed hypothesis is always to be preferred to an unconfirmed one. Moreover, as we saw, in his eagerness to invalidate inductive reasoning, he committed one of the most elementary errors of deductive reasoning!

#### 4. How far up the scale?

Underlying Goodman's riddle is another important question for inductive logic: how far up any scale of classification can generalizations legitimately be taken? Having for a given subject generalized a certain predicate, why not generalize further up the scale to a larger predicate?<sup>64</sup>

Consider a subject X and any two predicates S and G, related as *species and genus*, i.e. such that all S are G but not all G are S (i.e. some nonS are also G). Here, note well, S and G are both ordinary concepts, like green and colored.

---

contradictories in non-informative, he included contraries, viz. 'A or B' (where B is not A). This slightly conceals the issue, but does not in fact change it.

<sup>63</sup> See my *Future Logic*, chapter 50.

<sup>64</sup> I have touched upon this topic (indirectly, with regard to ethical logic) in my *Judaic Logic*, chapter 13.3.

- If all cases of X that we have observed so far are found to be S, and we have looked out for and not encountered any X that are not S, we may inductively infer that All X are S. This generalization remains valid so long as no cases of X that are *not* S are found; but if any X-nonS eventually do appear, we are required by inductive logic to revise our previous judgment, and particularize it to Some X are S and some X are not S. For induction proceeds conditionally<sup>65</sup>.
- The same reasoning applies to G<sup>66</sup>. Alternatively, granting that All S are G, we can from All X are S *deductively* infer that All X are G, by syllogism (1<sup>st</sup> Figure, AAA). Thus, we might postulate, if we are justified to generalize, for a given subject X, as far as the specific predicate S, we are also justified to do so higher still on the scale of classification, as far as the more general predicate G. This is logically okay if properly understood and applied.
- However, it would be a gross error of judgment<sup>67</sup> to *infer* from such valid generalization that there might be some X that are G but not S (even if we know there are *things other than X* that are G but not S). At this stage, the actual content All X are G is identical (in extension and implicitly in intension) to the All X are S from which it was

---

<sup>65</sup> For adduction or generalization is justified by *two* essential principles: (1) confirmation of a hypothesis by a positive instance, and (2) the non-rejection of the same hypothesis by any negative instance, and *both* principles must be equally obeyed for it to proceed logically. There are of course many other conditions involved – see my essay “Principles of Adduction” in *Phenomenology* (chapter VII.1).

<sup>66</sup> That is, given Some X are G (or deducing this from Some X are S), we can generalize to All X are G, provided there is no known negative instance (X-nonG) to belie it.

<sup>67</sup> This is as we saw one of the errors Goodman committed in formulating his “riddle”. This error is of a deductive rather than inductive nature.



derived<sup>68</sup>. How the two statements differ is only with regard to eventual corrective *particularization*...

- Suppose tomorrow we discover *an X that though still G is not S* (for example, an emerald of some color other than green). In such event, we would have to particularize the first (more specific) statement to 'Some X are S and some X are not S'; but the second (more generic) statement 'All X are G' would remain unchanged.<sup>69</sup>

- But as a result of such particularization All X are G has a *vaguer meaning*, since G no longer for us refers only to the S species of G but equally to some other (nonS) species of it. Thus, though the inductive rule would be to generalize as far up the scale as we indeed can go, we must keep in mind that the further up the scale we go, the more we dilute the eventual significance of our generalization.<sup>70</sup>

Thus, although in principle generalization up the scale is unfettered, in practice we proceed relatively slowly so as to maintain the noetic utility of our ideas and statements. To give a formal example: the proposition All X are S might be used as minor premise in a syllogism where S is the middle term, whereas the proposition All X are G – even if

---

<sup>68</sup> This is obvious if we consider that we may equally well obtain All X are G: (a) by generalization from Some X are G, which we deduce from Some X are S, or (b) by deduction from All X are S, which we generalize from Some X are S. In truth, it could be argued that these two are slightly different, since (a) requires that we make sure that there are no instances of X that are not G, whereas (b) requires that we make sure that there are no instances of X that are not S. This difference is however brought out in the ensuing stage of eventual particularization.

<sup>69</sup> Note that if we discover an X that is not G, it is necessarily also not S, given All S are G. In that event, both general propositions would of course have to be particularized.

<sup>70</sup> In this context, we could compare Goodman's "grue" concept to Feynman's concept of "oomph". The latter, defined (tongue-in-cheek) as "a kind of tendency for movement" might seem useful to "explain" various phenomena, but it is so vague that it cannot predict anything and is therefore worthless (p. 19).

still identical in extension and intension to the preceding – would be useless in that same context (i.e. with S as middle term).

Moreover, to regard All X are G as a more profitable generalization than All X are S, in the sense of providing us with information about more things for the same price in terms of given data, signals a confusion<sup>71</sup> between generalization for a given subject from a narrower predicate to a wider predicate, and generalization of a given predicate from a narrower subject to a wider subject.

The latter case is the truly profitable form of generalization. Suppose All X are P, and Y is an overclass of X (i.e. All X are Y, though not all Y are X), then this would consist in inducing that All Y are P — of course, unless or until some Y that is not P is discovered. The rules of such generalization are dealt with fully in my work *Future Logic* under the heading of Factorial Induction (Part VI).

---

<sup>71</sup> Which Goodman was guilty of in formulating his “riddle”, incidentally.

## 12. The Sorites paradox

*From Topics (published only online so far).*

### 1. What's a heap?

The Sorites paradox is not a paradox, in the strict sense of the term, but a question. The question is sometimes put in a sophisticated manner, so as to make it seem paradoxical. But it can be put in a more straightforward manner, in which case it is seen to be simple though not without importance. The term *sorites* is Latin, derived from the Greek *sōros*, meaning heap.

One way to express the Sorites paradox is: What is a 'heap' (of pebbles, say)? Or, how many pebbles (say) constitute a 'heap' of them? The obvious answer is there must surely be at least one pebble. If you have no pebbles, you do not have a heap, but a non-heap. But is one pebble enough? The obvious answer is: no, you need *at least two* pebbles to make a heap, since heap is a collective term, and one that additionally suggests that the pebbles are stacked one on top of the other (and you cannot stack a non-pebble on a pebble or a pebble on a non-pebble). A single pebble logically counts as a non-heap; heap is intrinsically plural. Formulated like that, the question is not very problematic. But if it is formulated as follows, it becomes more complicated. If we have many pebbles (say, 100) piled up, we obviously have a heap. What happens if we remove one pebble, do we still have a heap? Yes, 99 make up a heap. What happens if we remove one more pebble, do we still have a heap? Yes, 98 make a heap. And so on, till we come

to low numbers, at which stage the sophist wonders whether two pebbles constitute a heap, then one pebble, then no pebble. At this stage, the question appears paradoxical, rather artificially: should we conclude that no pebble makes a heap, or one pebble makes a heap? The sophist thinks we might; the real logician knows we cannot.

## 2. The use of vague terms

Clearly, the problem here, insofar as there is one, has to do with the exact formulation of initially vague terms. If we do not *at the outset* step back and think about the exact intent of a vague term, of this sort (a term suggesting quantity, as it happens) or any other, we may find ourselves in difficulty *further on* in our discourse. So, we need to stop and think before use of such terms, and preempt any difficulties they might eventually create. In the present case, as we have seen, the term (heap) is inherently plural (and so inapplicable to less than two pebbles). A sophist prefers to complicate the matter, so as to put human reason and knowledge in doubt, as is his wont; but the matter is really simple enough.

In some cases, to be sure, there is a *conventional* element to the definition of a vague term. This can be illustrated with reference to another version of the so-called Sorites paradox: As of how many hairs is a man not bald? Obviously, a man with no hair at all is definitely bald. But would a man with only one hair, or a very small number of hairs, not be considered bald in ordinary discourse? Perhaps so – but only conventionally. If we define bald *strictly*, it implies zero hair; but if we define it *loosely*, it may include cases involving an *arbitrary*, though preferably small, number of hairs, determined by

convention – for examples, five or twenty hairs<sup>1</sup>. Clearly, if we want to avoid confusion, nothing stops us from referring to this broadened sense of ‘bald’ as, more accurately, ‘bald or almost bald’.

This issue of vagueness is nothing special – it is not limited to terms giving rise to the Sorites paradox. For instance, a relative term like ‘small’ (or its relative, ‘large’) is *inherently vague* and can only be used with precision in specific situations by means of a conventional quantity. Again, when dealing with continua, we may need to set arbitrary dividing lines. For instance, there is no objective dividing line between one color and an adjacent color in the spectrum, and it may be necessary in some circumstances for us to imagine one (e.g. for legal or other practical purposes<sup>2</sup>). Conventional distinctions are part of human thought; but, it is important to stress, they are not all of human thought. There are always objective elements behind conventional ones. For example, the dividing line between blue and green would be somewhere in between what we see as clearly blue and what we see as clearly green – it would never be far on one side or the other, and much less between green and yellow or between blue and indigo or still further afield.

---

<sup>1</sup> Or whatever minimal hairiness seems to us subjectively as so close to bald as to be effectively bald.

<sup>2</sup> For example, in Jewish law (*halakhah*) much attention is given to quantitative definitions, notably to the maxima or minima of durations, times o’clock, distances, lengths, volumes, weights, temperatures, monetary values, etc. Initially, such measures were often expressed by the rabbis in vague terms (e.g. ‘the volume of an egg’), but later more precise formulations were called for (which different authorities might differently estimate). However, some measures remain subjective (e.g. the estimate of when one is full after eating). See for more details: [http://halachipedia.com/index.php?title=Reference\\_of\\_Measurements\\_in\\_Halacha](http://halachipedia.com/index.php?title=Reference_of_Measurements_in_Halacha).

Returning to our alleged paradox, a few more comments are in order. As already stated, there is nothing paradoxical in the concept of a heap, if it is properly defined. Most simply, a heap can be defined as material items placed on top of each other in whatever way, implying that there must be two or more items. A more complex definition would assume that a heap must be pyramidal, i.e. requires at least four such items (three for the base and one on top); but this puts us in no difficulty, as it can be referred to more specifically as a pyramidal heap. Similarly, baldness *stricto sensu* refers to no hair at all; it is nevertheless applied to small quantities of hair, though only roughly-speaking.

There are many vague terms of this sort in our common discourse<sup>3</sup>, some of which may require a conventional definition for pragmatic reasons. For example, the term ‘crowd’ might be taken to refer to a gathering of three or more people, on the basis of the popular saying that “two’s company, but three’s a crowd;” or we might, say in software used by the police to monitor large groups of people, opt for a larger minimum (say, 50 or 500), set arbitrarily as cause for alarm. The word ‘mob’ might be preferred when the latter crowd goes on a rampage.

But in any case, there is no real logical problem in such unspecific quantitative expressions. They do not constitute a defect in ordinary language, requiring us to construct an

---

<sup>3</sup> Indeed, this is inevitable on two counts. First, many concrete objects are impossible to precisely define. Where, for example, does an orange end precisely? Is the perfume or heat emanating from the fruit part of it or not? At what points in time and space may such emanations logically be regarded as separate from it? Second, human knowledge being inductive, we cannot always start a concept with a precise definition, but tend to leave it open, to be defined more and more precisely as experience unfolds. In this perspective, the majority of abstract terms we use are open, including terms that may be used to more precisely define other terms – so, here again, vagueness is inevitable. But such ontic and epistemic difficulties do not imply paradox; they simply call for philosophical reflection.

“ideal language” where all terms have single precise meanings<sup>4</sup>. Much less do they call for treatment by means of abstruse symbolic logics (dearly loved by many modern logicians). On the contrary, they demonstrate the versatility and flexibility of ordinary thought and speech; and they witness the fact that much of our linguistic communication has non-verbal undercurrents, which we mostly comprehend very ably. Most of our daily use of vague terms involves *no need* for more clear-cut definition. They are used to suggest things approximately, and are not intended as precise and true affirmations. If the need for precision and truth does arise, it is then addressed in a way that preserves consistency (by explicit convention if necessary).

### 3. Reasoning with vague terms

To be sure, vague terms can be perilous if we try to reason with them. But vague terms are often used without involving them in argumentative processes. Moreover, reasoning with vague terms is not always invalid – there are contexts where the vagueness does not inhibit a reliable conclusion.

In categorical (or hypothetical) syllogism, the rule regarding vague terms (or theses) is the following: The

---

<sup>4</sup> The resort to an “ideal language” by certain modern logicians to solve a problem of logic is futile. Unable to understand the actual way we real human beings logically deal with certain cognitive difficulties, they try to impose a superficial, artificial and impractical way of thinking on the rest of us. The role of the genuine logician is not to impose imaginary logics, but to understand our natural logical means and thence to perfect and reinforce them. Reasoning by humans should be the central concern of logicians. The natural language way to deal with Sorites paradoxes is to use words more precisely – e.g. instead of calling persons with very few hairs ‘bald’, to call them ‘almost bald’; or more accurately still, if necessary for some practical purpose, ‘having (say) one to ten hairs’.

middle item (term or thesis, as the case may be) cannot be vague, because it would provide no guarantee that its intent is the same in the two premises. If the middle item is vague, we cannot be sure of overlap and the conclusion is invalid. On the other hand, the major and minor items can be vague without affecting the argument, and that in all four figures. There is, however, an exception to this rule – when the middle term is vague, but not so vague that overlap is not guaranteed, a valid argument can still be made.

The latter is evident when we consider the following two arguments in the third figure, in which the middle term ensures overlap even though neither premise is universal. The expressions ‘most’ (more than half) and ‘few’ (less than half) are vague, insofar as they do not specify exact numbers. But notice the particularity (as against majority or minority) in the conclusion – i.e. the increased vagueness of the conclusion.

Most M are P  
 and Most M are S;  
 therefore, Some S are P

Few M are P (which implies that Most M are not P)  
 and Most M are S;  
 therefore, Some S are not P

In apodosis, the rules regarding vague terms or theses are the following:

Modus ponens:

If A is B, then C is D, and A is B (affirming the antecedent);  
 then, C is D (consequent is affirmed).



The antecedent and the minor premise cannot be vague; else, the conclusion is invalid. However, the consequent could be vague, and the conclusion would still be valid (though also vague).

Modus tollens:

If A is B, then C is D, and C is not D (denying the consequent); then, A is not B (antecedent is denied).

The consequent and the minor premise cannot be vague; else, the conclusion is invalid. However, the antecedent could be vague, and the conclusion would still be valid (though also vague).

Here again, in both moods, exception is conceivable, if we know that the major and minor premises overlap, even if we don't know precisely how much they overlap.

Similar rules may be formulated for other varieties of argument, such as dilemma or a fortiori.

The people who claim that vague terms are inherently paradoxical are dishonestly nitpicking, motivated by the desire to impress themselves or others by their ability to find and resolve (contrived) paradoxes, or (worse still) to demonstrate that human knowledge is inevitably paradoxical and therefore futile. Clearly, just as it is dishonest to call a single pebble or no pebble a heap, it is dishonest to call a person with one or more hairs bald<sup>5</sup>. If

---

<sup>5</sup> As I explain in *A Fortiori Logic*, chapters 1.4 and 2.5, it is sometimes useful to formulate terms in a way so inclusive that positive, zero and negative values are all embraced by them. This is often done in scientific discourse because it facilitates some calculations and graphs. But it must be well understood that such inclusive terms are inherently undeniable – i.e. they already englobe both an affirmation and its denial. In the present context, we might choose to enlarge terms like heap or bald to include their opposites, for whatever reason; but when we do so we must remain keenly aware of what we are doing. If

you indulge in such contradictions-in-terms to start with, you are bound to end up with paradoxes. People who behave thus are not real logicians but sophists. They spin and fabricate – they are not interested in finding ways to true knowledge.

#### 4. Making up fake paradoxes

The original formulations of both the conundrums described above, relating to a heap and to baldness, are attributed to Eubulides of Miletus (*fl.* 4<sup>th</sup> Cent. BCE), the Megarian logician who also gave us the Liar paradox<sup>6</sup>. He was a student of Euclid, a teacher of Diodorus Cronus, and a contemporary and rival of Aristotle. These puzzles were perhaps not initially presented as paradoxes, but rather as illustrations of a question (*viz.* where should we draw the line?). This possibly reflected a dawning consciousness that there are vague terms that may require arbitrary definition in some circumstances. As above shown, this problem is solved easily enough. However, later thinkers tried to make a mountain out of a molehill, and presented the issue in the form of an argument-chain (or sorites, where the conclusion of the preceding argument serves as a premise for the next).

Thus, the bald man puzzle became, in its positive formulation: Surely, a man with one hair is about as bald as one with no hair; and if a man with only one hair can

---

we do not treat such terms with appropriate care, we should not be surprised if we are forced into contradictions.

<sup>6</sup> A couple of centuries earlier, Epimenides of Knossos declared: “Cretans, always liars,” though himself a Cretan, apparently unaware of the contradiction inherent such a statement. Eubulides may have noticed the paradox involved and sought to refine it and strengthen it, since it was not a double one but one easily resolved by saying that possibly not all Cretans are liars or that Cretans do not all always lie (Epimenides being a notable exception).

still be called bald, then a man with two hairs qualifies as bald; and if a man with two hairs is bald, then a man with three hairs is bald; etc.; therefore, a man with a thousand hairs can still be considered bald (paradox). Alternatively, the argument could be stated in negative form: if a man with a thousand hairs is not to be regarded as bald, then one with 999 hairs is not to be so regarded either; and if 999 does not qualify, then 998 does not either; and so on... whence, a man with one hair only is not bald; therefore, a man with no hair is not bald (paradox).

Clearly, these arguments are *forced* – they involve some very doubtful and misleading premises. In the positive version, the false premise is that ‘a man with no hair can be called bald’ *implies* ‘a man with one hair can be called bald’. In the negative version, the false premise is that ‘a man with one hair cannot be called bald’ *implies* ‘a man with no hair cannot be called bald’. It is the same false claim of implication (in contrapositive form). The paradox is created by the refusal to admit that ‘bald’, strictly-speaking, means ‘hairless’; which refusal is not based on honest logical insight, but on a willful act of illogic. Similarly, in the case of a heap, the trick consists in implying that a single pebble or even no pebble may be considered as a heap.

The inventors of these arguments do not pause and rationally reflect on the underlying issue (i.e. where is the dividing line?) before engaging in an apparent inference process, but instead attempt to bamboozle us into a paradoxical corner. The argument-chain proposed just serves as a smokescreen to conceal the crucial false claim being put over. They are the logical equivalents of pyramid sales, each sale supporting the next without solid foundation. People who are taken in by the tricky move are simply bad logicians, if not shamelessly dishonest. They then pretentiously weave massive and intricate theories

around this phenomenon, untroubled by the initial error or lie in their discourse.

To those who argue that a single hair or pebble hardly makes any difference, I would suggest that they make the following simple physical experiment: take an accurate balance with the same weight on both sides, then add a single hair or pebble to one side and watch the scales tip! To those still unconvinced by this, because they dogmatically believe that logic is a matter of fancy and convention, I would suggest (tongue-in-cheek, more or less) that they place, under the heavier scale, a plunger connected to an explosive device strapped to their nose, and then watch Reality blow up in their face! That argument, I think, might finally convince them, if they survived.

As regards Eubulides, we might note in passing his other paradoxes. The most significant is of course the Liar paradox, which as I have shown in detail elsewhere<sup>7</sup> is exceptionally powerful due to the variety of difficulties it involves (but still quite resolvable). Another three paradoxes<sup>8</sup> deal with equivocations in the term ‘know’, specifically with failure to immediately recognize someone one normally recognizes immediately (such as a close relative or old friend), when the latter is masked or has been away too long or is not looked at attentively enough. Another, the Horns paradox claims that what you have not lost must be in your possession; whence, if you have not lost horns, you must have horns. This apodosis involves a false major premise, since something one has ‘never had’ may equally (as well as something one ‘still has’) be characterized as ‘not lost’ – so the consequent does not follow upon the antecedent; therefore, if one has not lost

---

<sup>7</sup> See my *A Fortiori Logic*, Appendix 7.4.

<sup>8</sup> Namely, the Masked Man, Elektra and Overlooked Man paradoxes.

horns, one cannot be assumed to have horns (since one may well never have had any).

From this short list, it can be seen that Eubulides' queries all give rise to some sort of reflection on logic – reflections on vague terms, on conventional definitions, on equivocations, on term-negations, on self-reference, and various other difficulties that may arise in human discourse. It would be wrong, I think, to assume the motive of such queries to have been teasing or obfuscation (although, to be sure, later skeptics did use such conundrums malignantly, as already mentioned). Rather, I'd say, they served as springboards for earnest reflections and discussions on logic – because it seems unlikely that they were formulated without any attempt to solve the problems they engendered. In some cases, valid explanations or resolutions were no doubt proposed (even though they may not have come down to us), while in other cases the difficulties may have seemed insurmountable. In other words, I doubt that Eubulides was merely sophistical – I'd class him rather as a serious logician.

## 13. Protagoras vs. Euathlus

*From Topics (published only online so far).*

### 1. An ancient paradox

One of the many alleged paradoxes that have come down to us from the Greeks is the dispute between Protagoras (of Abdera, ca. 480-410 BCE) and his student Euathlus (about whom nothing more is known). The story is told by Aulus Gellius (Roman, ca. 125-180 CE)<sup>80</sup>, that Protagoras, a famous Sophist, and an expensive teacher, agreed with Euathlus to train him in rhetoric, a discipline essential at the time to argumentation in courts of law. The agreement was that Euathlus would not have to pay Protagoras the specified fee (or the unpaid portion of the fee, by some accounts) until he had been fully trained and went on to plead his first case and win it<sup>81</sup>.

It is said that after Euathlus completed his course, he did not (for whatever reason) choose to use his newly acquired skills before any court of law, and so he never won or lost any case, and so was contractually not required to pay Protagoras anything. Nevertheless, Protagoras, with motives that we shall presently consider, sued him (in the court of Areopagus in Athens). Euathlus chose to

---

<sup>80</sup> Some 550-600 years after the fact, in his *Attic Nights* (ca. 150 C.E.).

<sup>81</sup> The contract in question was presumably verbal, rather than written, in those days; but we may take it that both parties agreed on its stated clause.

personally defend himself. The following arguments were reportedly put before the court:

- a. **Protagoras** argued that he surely ought to be paid the fee, because (i) if the court ruled in his favor, he could on that basis demand payment; and (ii) if the court ruled in favor of the defendant, then the latter would have won his first case and therefore be contractually obliged to pay the fee anyway.
- b. **Euathlus** replied that he surely ought *not* to pay the fee, because (i) if the court ruled in favor of the plaintiff, then Euathlus would have lost his first case and therefore not be contractually obliged to pay the fee; and (ii) if the court ruled in favor of the defendant, then he would on that basis be exempt from payment anyway.

Thus, while the plaintiff argued, apparently convincingly, that he was certain to deserve payment however the court ruled, the defendant was in turn able to argue, apparently just as convincingly, that he was certain to be exempt from payment in either event. For this reason, this case is regarded as paradoxical. It is said that the court was so confused by these arguments and counter-arguments that it chose to adjourn *sine die* (or, some say, for 100 years) to avoid judgment.

The significance of this legal dispute for logic and philosophy is that it gives the impression that two people can argue dilemmatically and paradoxically in opposite directions and both be right. The enemies of human reason relish this kind of conclusion, since it would put in doubt the reliability and efficacy of human reason. But as we shall now show, the said impression is very superficial. There are, to my mind, at least two possible resolutions of this so-called Paradox of the Court. In fact, although I thought of the second before the first, the latter logically precedes the former. I later learned from Peter Suber's survey of the

literature on the subject<sup>82</sup> that the first resolution was long before me proposed by Aulus Gellius; the second resolution seems to be novel.

## 2. First resolution

The simplest solution to this problem is to suppose that the wily Protagoras, seeing that Euathlus was taking his time getting to work, decided to speed things up. Protagoras trapped his pupil by using the above argument, knowing full well that he would lose a first trial, but win an eventual second trial. He knew he would lose a first trial, because the agreement between the parties only obligated Euathlus to pay the fee once he won his first case; it did not obligate him to practice law anytime soon, or even ever. Euathlus foolishly fell into the trap and personally argued the case in court. Had the court not adjourned *sine die*, it would have logically given him victory, thus making Euathlus win his first case. Thereafter, assuming that a second trial was legally permitted – and both the parties’ arguments above do make this assumption – Protagoras would have been the ultimate winner. Of course, no second trial would be necessary if Euathlus conceded that having won the first trial he was sure to lose the second, and settled the account forthwith.

In other words, Protagoras’s first argument (i) was mere camouflage; he was really relying on his second argument (ii). Euathlus let his vanity get the better of him and formulated two fancy counter-arguments, thinking to outdo his teacher. But Protagoras was more cunning than him.

---

<sup>82</sup> In *The Paradox of Self-Amendment*. The section on the Protagoras v. Euathlus paradox can be read online; it is worth reading, including the notes, and not very long. See at: <https://dash.harvard.edu/bitstream/handle/1/10288413/Peter%20Suber%2c%20Paradox%20of%20Self-Amendment%2c%20Section%2020.html?sequence=1#A>.



The only way Euathlus could have avoided being beaten was by hiring a lawyer<sup>83</sup>. If the lawyer won the first trial, as he could be logically expected to since the only condition for defeat here was not satisfied, Euathlus would not be considered as having personally won his first case in court, and Protagoras would not be able to win a second trial. Thus, the master was indeed superior to the pupil. There was no real paradox, since there was an actual way out of the apparent paradox.

Of course, one might add that Euathlus was in practice the winner, since through his counter-argument he managed to so bewilder the court that it gave up trying to judge the matter at all, and he was not forced to pay up. Maybe he hoped for that and he lucked out. But on a theoretical, logical level, in the present perspective, he proved to be not too intelligent. Not only did he foolishly not hire a lawyer to plead on his behalf, but he also wrongly assumed that his argumentation was effective in countering Protagoras'. He kidded himself into thinking that if he won the first trial, he would not have to face a second one. He should have examined his teacher's argumentation more carefully.

Let us now look at the arguments in more formal terms, to clarify them. I shall introduce the following symbols: let P = Protagoras, Q = Euathlus; A = the agreed condition that Q wins his first case in order to be liable, and C = the agreed result that fee must be paid by Q to P. We know at the outset that if A then C, and if not-A then not-C: these are the terms of the agreement. The arguments are as follows (with my critical commentary in *italics*):

- a. According to P: (i) if court rules that P wins, then C is true (*but objectively court cannot rule for P, since A not yet true, so this is a non-starter*); whereas if (ii) court

---

<sup>83</sup> Assuming this was possible in Athens in those days. This is a fair assumption (even if some commentators deny it) since, after all, Euathlus was apparently trained by Protagoras to be a lawyer himself.

rules that Q wins (*as it logically must*), then not-C is true (*at the conclusion of a first trial*); but if Q thus wins, then A becomes true and C must (*in a later trial*) follow, in which event P finally wins.

- b. According to Q: (i) if court rules that P wins, then not-A is true (*at the conclusion of a first trial*); but if not-A is true then not-C must (*in a later trial*) follow, so Q finally wins (*but objectively court cannot rule for P, since A not yet true, so this is a non-starter*); whereas if (ii) court rules that Q wins (*as it logically must*), then not-C is true (*but here Q fails to mention later consequences that P rightly pointed out*).

In conclusion, P is logically the resultant winner; Q's arguments are in fact insufficient to prevent this outcome. P pretends to seek to win immediately; but in truth his aim is longer term victory (in the second round). Q imagines he might lose the first round but win the second or that he might win the first round without having to face a second; but these are all fantasies. It is difficult to understand why the court found this case too confusing – the judge (or judges) can't have been very bright fellows.

### 3. Second resolution

A more complex solution to the problem is as follows. It is possible that Protagoras sued Euathlus by appealing to *an unspoken clause* of the agreement. The agreement contained *only one explicit clause*, viz. that Euathlus would have to pay Protagoras the fee if and only if he won his first case. If that was so, Protagoras would have no basis for requesting a trial, since that condition had obviously not been satisfied. But since he sued, he must have thought and argued that the agreement included a *tacit (or perhaps implicit) understanding* that Euathlus would practice law

within a reasonable time lapse<sup>84</sup>, at which time his new skills would be tested in a court and he would be expected to pay if he won. Protagoras couldn't have imagined the judge would allow a trial to proceed, let alone would rule in his favor, without some good reason<sup>85</sup>.

Clearly, what the judge was called upon to decide in this trial was (could only have been, in the present perspective) whether this claim by Protagoras, that there was a tacit clause to the agreement, was justified. He could well have justified it by considering that had Euathlus been allowed not to practice law at all or to practice it as late as he chose, the agreement would have surely specified the caveat. He could equally well have rejected it by considering that Protagoras took for granted something he should have explicitly obtained agreement on. So, the case hinges on a tacit issue, rather than exclusively on the explicit clause of the agreement; i.e. there was more to the story than is told.

Furthermore, it is evident from the arguments presented by both parties in this trial that each of them foresaw the possibility of a second trial in which the ruling of the first trial could be reversed. This is logically implied in the second argument of Protagoras and in the first argument of Euathlus. In these two eventualities, a second trial would be needed to finalize the judgment; unless, of course, the losing party freely concedes its inevitable result and settles the account in advance. The judge in the first trial could not decide in favor of either party and then against him in the same breath. A first judgment of win or loss would have to be established before a second judgment could be made, in

---

<sup>84</sup> Protagoras would also, of course, claim that the reasonable delay had expired. If the court agreed with the existence of a tacit clause but disagreed with the claim it was fulfilled, that would merely adjourn the case for a certain amount of time (of their estimate).

<sup>85</sup> Needless to say, the present analysis is made entirely from a logical viewpoint, although a court of law might reason differently or even not reason at all (e.g. bribery, favoritism, pressure).

view of the terms of the contract. So, there is a time factor to take into consideration in analyzing these arguments.

It should be noted that, whereas the first trial has an uncertain outcome, since it depends on the decision (the judgment call) of the judge regarding an alleged tacit clause to the agreement, the second trial has a foreseeable outcome, since it depends solely on the explicit clause of the agreement.

Clearly, as we shall now show, if we factor the above considerations into the arguments, the paradoxical appearance is easily dissolved. I shall here use the following symbols: let P = Protagoras; Q = Euathlus; A = the explicit condition that Q wins his first case; B = the alleged implicit condition that Q was required to practice within a reasonable amount of time; and C = the fee must be paid by Q to P. The arguments are as follows:

- a. According to P: (i) if B then C; and (ii) if not-B then not-C, *but if not-C then (later) if A then C*; therefore, C anyway.
- b. According to Q: (i) if B, then C, *but if C then (later) if not-A then not-C*; and (ii) if not-B, then not-C; therefore, not-C anyway.

From this we see that the two parties' arguments are much alike, but each side has cunningly left out part of the consequences (shown in *italics*). P has truncated the consequences of (i) that Q points out, and Q has truncated the consequences of (ii) that P points out. For this reason, they seem to balance each other out. But if we take *all* the subsequent events (in a possible second trial) into account, we get the following more objective joint argument (c), which is clearly non-paradoxical:

- (i) if B then C, but if C then (later) if not-A then not-C; and
- (ii) if not-B then not-C, but if not-C then (later) if A then C.

So, in (i) the final conclusion is not-C, while in (ii) it is C – which means that there is no paradox. This also shows that, while it cannot objectively be predicted whether P or Q will win the first trial, it can be said that (i) if P wins the first trial, he will lose the second, and if Q wins the first trial, he will lose the second. Obviously, P cannot argue that he has a right to a second trial (if he loses the first) but Q has not; likewise, Q cannot argue that he has a right to a second trial (if he loses the first) but P has not. So, we must take all later events into consideration to logically reconcile all the arguments. Note that if for some reason there is no second trial, there is also no paradox, since the conclusion will be either (i) C or (ii) not-C.

Whatever happens, there is no paradox because neither party can in fact, contrary to initial appearances, claim inevitable victory; victory does not come both ways for both parties, but only one way for each party.

Clearly, here, both parties were employing the common sophistical trick of hiding an inconvenient part of the unbiased argument from the court. Euathlus was a good apprentice of Protagoras', since his counter-argument exactly mirrors the latter's argument. That is, there was an element of dishonesty in both their arguments; both were intellectual frauds at heart, knowingly expounding half-truths. Therefore, this fake paradox presents no deep challenge to Logic, contrary to the claims of Relativists. In particular, Protagoras' general claim that "there are two sides to every issue" (*duo logoi*) is shown to be spurious in the present context.

It is worth always keeping in mind that some people involved in philosophy and logic, as in life in general, are sometimes moved by the evil impulse; indeed, some much more than others. They may consciously lie, or subconsciously twist facts and arguments, for a large variety of motives. Usually, lusts for power, fame and

fortune play some role. An academic may want the admiration of his peers or of his students; a husband may want to impress his wife; an unemployed may hope to get a job; and so on. It is wrong to look upon all philosophical statements as disinterested. Philosophers and logicians are not all pure scientists or saints.

I might add that the secret of success with finding solutions to philosophical and logical problems, and particularly to paradoxes, is the sincere desire to do so. Many philosophers and logicians approach problems with a negative attitude, not really wanting to solve them, but rather wishing to rationalize their antipathy to human reason through them. The honest researcher is moved by his better impulses; he is sincerely desirous to confirm the effectiveness of the human faculties of cognition – that's precisely why he succeeds.

#### **4. Inadequate resolutions**

As earlier mentioned, based on Suber's account of the literature, the first resolution should be attributed to Aulus Gellius, but the second resolution seems to be original. Suber's account shows that the court paradox has been discussed in a number of works over time, but more often apparently from a legal point of view than from a logical one. The legal issues involved are manifold, but most need not and should not be taken into consideration in a purely logical perspective. Why? Because the logician's purpose here is not to decide the case, i.e. who should win or lose, but merely to explain and remove the appearance of paradox.

This remark can be illustrated with reference to the resolution (not mentioned by Suber) proposed centuries ago by Lorenzo Valla (Italy, ca. 1406-57). This attempt at resolution is not adequate because it relies on a thoughtless distinction between payment on account of the court's

verdict and that on account of the agreement. I quote an SEP article which describes it<sup>86</sup>:

“If Euathlus loses the case, he will have to pay the rest of the fee, on account of the verdict of the judges; but if Euathlus wins, he will also have to pay, this time on account of his agreement with Protagoras. Euathlus, however, cleverly converts the argument: in neither case will he have to pay, on account of the court's decision (if he wins), or on account of the agreement with Protagoras (if he loses)... Briefly put, Valla says that Euathlus cannot have it both ways and must choose one or the other alternative: he must comply either with his agreement with Protagoras or with the verdict passed by the judges. If they decide against Protagoras, he may try to reclaim his money in a second lawsuit.”

Let us analyze Valla's proposed resolution using the following symbols: let P = Protagoras; Q = Euathlus; and C = the fee must be paid by Q to P.

- a. According to P: If Q loses (for whatever reason), then C (by verdict); but if Q wins, then C (by agreement).
- b. According to Q: If Q wins (for whatever reason), then not-C (by verdict); but if Q loses, then not-C (by agreement).

Valla's conclusion, as here presented<sup>87</sup>, is unclear. Apparently, he puts the onus on Euathlus in particular to

---

<sup>86</sup> See *Stanford Enc. of Phil.*, online at <https://plato.stanford.edu/entries/lorenzo-valla/>. This issue is apparently treated in *Repastinatio dialectice et philosophie*. The author of the article is apparently Lodi Nauta (2013).

<sup>87</sup> I have not read Valla's work. I have to assume that the author of the SEP article on Valla correctly presented Valla's reasoning. I

“choose one or the other alternative” and comply with either the agreement or the verdict. This tells us nothing, since Euathlus’ argument shows he is willing to comply with either, except that he projects both as in his favor. Valla adds that if Protagoras loses a first lawsuit, he may win a second. But here again, this does not resolve the paradox, but only repeats one part of it.

From this we see that Valla has not thought the issues through: he does not consider on what grounds Q might lose or win ‘by the court’s verdict’; and he does not realize that the second argument by each party, where Q might alternatively win or lose ‘by the agreement’, in fact refers to a second lawsuit, since the agreement is evidently not about to be implemented voluntarily (Valla’s mention of a second lawsuit is placed beyond the four if-then statements, which themselves do not emphasize the temporal sequences involved). In the last analysis, then, Valla does not arrive at a ready resolution of the paradox.

The trouble with Valla’s approach is that it effectively takes the initial decision of the court to be arbitrary, i.e. unrelated to the agreement between the parties. It does not consider on what basis the court might judge that Euathlus can lose the first round. This may be acceptable legally, but it is not acceptable logically – and our concern here is with logic. Logically, the idea that Euathlus can lose the first round is a non-starter, if we go by the explicit clause of the agreement. He might lose the first round only if the court grants the supposition that there was a tacit clause to the

---

suspect he or she may not have, as the details given in the article are rather vague and inconclusive; they do not clarify exactly what resolution of the paradox Valla had in mind. Not everyone is good at logic. (Note that I did write to the author, asking him or her to please verify the summary of Valla’s view given in the SEP article, but I got no reply. That the author did not deign to respond confirmed my estimate that this is not a very reliable source: intelligent people confidently welcome reasonable queries.)



agreement, such as that he had to practice law sometime soon.

In any case, the verdict for the first round cannot be arbitrary – i.e. irrational, unjustified - but must directly relate to the agreement. This is equally true for the second round. So, no disconnect between verdict and agreement is logically permissible.

Interestingly, my first reaction to the paradox a few years ago<sup>88</sup> was very similar to Valla's. But as soon as I set about seriously considering the issues for the present essay, I realized its weakness and uselessness. The paradox appears neutralized if we insert a distinction between 'payment following court verdict' and 'payment following contract terms' and assign different symbols to these two consequences, say C1 and C2. These two terms may or may not be quantitatively identical; but they anyway refer to distinct events. This measure seems to nullify the paradox, because the consequences of the first and second if-then statements would be different for each party. Thus:

- a. According to P: If P wins, then C1; but if P loses, then C2.
- b. According to Q: If Q wins, then not-C1; but if Q loses, then not-C2.

It follows that if P wins and Q loses, then C1 and not-C2 are true; and if P loses and Q wins, then not-C1 and C2 are true. Since C1 and C2 are different terms, even if they happen to refer to the same monetary amount, the consequents in each party's argument are not formally contradictory (since the defining motive is different), so there is no paradox. But, to repeat, this approach does not explain why the court would give a verdict inconsistent with the agreement, so it is artificial. For a genuinely logical resolution, we must focus attention on the

---

<sup>88</sup>

In 2009, in an e-mail to someone.

agreement, and not admit arbitrary verdicts. Arbitrary verdicts just muddy the waters.

Other attempted resolutions I have seen are also flawed, either because they similarly conflate legal and logical issues, or because they do not follow through on all the consequences of all the suppositions. In any case, either of my proposed two resolutions suffices to unravel the paradox; but the two together take care of all eventualities. Note that, in the first resolution, we not only resolve the paradox, but also incidentally decide the case (in favor of Protagoras); whereas in the second resolution, we are content as logicians to resolve the paradox, leaving the task of legal decision to the judge.

To conclude: the paradox of the court is due to a number of factors, which must be untangled and taken into consideration if it is to be resolved. (a) The terms ‘win/lose’ cannot refer to arbitrary judgments by the court; if they do, the paradox may be perpetuated. (b) It must be realized that the arguments put forward by the two parties imply that the process of resolution has two phases: a first trial followed by a second trial (or by a ready concession and settlement without need of a second trial); the appearance of paradox is partly due to overlooking this time factor. (c) There is some vagueness in what is meant by ‘the agreement’, and all possible interpretations must be taken into account: if the agreement is taken to refer exclusively to the explicit clause, then the first trial concerns that only and is easily decided; but if the agreement is claimed to involve a tacit understanding, then the first trial aims at a decision regarding that tacit clause and the second round deals with the explicit clause.

## 14. Buddhist antinomic discourse

*From Buddhist Illogic 1 & 2.*

### 1. The tetralemma

Western philosophical and scientific thought is based on Aristotelian logic, whose founding principles are the three “Laws of Thought”. These can be briefly stated as “A is A” (Identity), “Nothing is both A and non-A” (Non-contradiction) and “Nothing is neither A nor non-A” (Exclusion of the Middle). These are not claimed as mere hypotheses, note well, but as incontrovertible premises of all rational human thought<sup>89</sup>.

Religions like Judaism, Christianity and Islam, even while adhering to these laws in much of their discourse and paying lip-service to them, in their bids to interpret their own sacred texts and to make their doctrines seem reasonable to their converts, have often ignored these same laws. This is especially true of mystical trends within these religions, but many examples could be given from mainstream writings. The same can be said of some aspects of Buddhist philosophy.

The *tetralemma*<sup>90</sup> is a derivative of the laws of thought, with reference to any two terms or propositions, labeled A

---

<sup>89</sup> See my *Future Logic*, ch. 2 and 20, and later essays on the subject (published on my website [www.the逻辑ian.net](http://www.the逻辑ian.net)).

<sup>90</sup> See Cheng, pp. 36-38, on this topic. He there refers to MT opening statement, as well as XVII:12a and XXIII:1a. Etym. Gk. *tetra*

and B, and their opposites non-A and non-B. Four combinations of these four terms are conceivable, namely “A and B” (both), “non-A and non-B” (neither), “A and non-B” and “non-A and B” (one or the other only). According to Aristotelian logic, these four statements are incompatible with each other (*only one of them can be true*, because if two or more were affirmed then “A and non-A” or “B and non-B” or both would be true, and the latter implications are self-contradictory) and exhaustive (*at least one of them must be true*, since if they were all denied then “not A and not non-A” or “not B and not non-B” or both would be true, and the latter implications go against the excluded middle).

Now, what Nagarjuna does is insert the term A in place of B (i.e. he takes the case of  $B = A$ ), and effectively claim that the above four logical possibilities of combination apply in that special case – so that “A and A (=B)”, “non-A and non-A (=non-B)”, “A and non-A (=non-B)”, “non-A and A (=B)” seem logically acceptable. He then goes on to argue that there are four existential possibilities: **affirmation of A ( $A + A = A$ )**, **denial of A ( $\text{non-A} + \text{non-A} = \text{non-A}$ )**, **both affirmation and denial of A (A and non-A)** and **neither affirmation nor denial of A (not A and not non-A)**. He is thus apparently using the principles and terminology of common logic to arrive at a very opposite result. This gives him and readers the impression that it is quite reasonable to both affirm and deny or to neither affirm nor deny.

But in Aristotelian logic, the latter two alternatives are at the outset excluded – “both A and non-A” by the Law of Non-contradiction and “neither A nor non-A” by the Law of the Excluded-Middle – and the only logical possibilities left are “A” or “non-A”. The anti-Aristotelian position may

---

= four, *lemma* = alternatives. Term coined in contrast to the dilemma “A or non-A”.

be viewed, in a positive light, as an anti-Nominalist position, reminding us that things are never quite what they seem or that things cannot be precisely classified or labeled. But ultimately, they intend the death of Logic; for without the laws of thought, how are we to distinguish between true and false judgments?

The law of identity “A is A” is a conviction that things have some identity (whatever it specifically be) rather than another, or than no identity at all. It is an affirmation that knowledge is ultimately possible, and a rejection of sheer relativism or obscurantism. Nagarjuna’s goal is to deny identity.

It should be noted here that Aristotle is very precise in his formulation of the law of contradiction, stating in his *Metaphysics* “The same attribute cannot *at the same time* belong and not belong *to the same subject in the same respect*” (italics mine). Thus, an alternative statement of the laws of thought would be the ‘trilemma’ (let us so call it) “*either wholly A, or wholly non-A, or both partly A and partly non-A*”, which excludes the fourth alternative “both wholly A and wholly non-A”. The Buddhist attack on the laws of thought draws some of its credibility from the fact that people subconsciously refer to this ‘trilemma’, thinking superficially that indeed opposite things may occur in the same place at different times or at the same time in different places or in various respects, without thereby giving rise to logical difficulty incapable of resolution. But it should be clear that the Buddhist position is much more radical than that, accepting thoroughgoing antinomy.

Similarly with regard to the law of the excluded middle, which affirms the situation “neither A nor non-A” to be impossible *in fact*. People are misled by the possibility of uncertainty *in knowledge*, as to whether A or non-A is the case in fact, into believing that this law of thought is open

to debate. But it must be understood that the thrust of this logical rule is inductive, rather than deductive; i.e. it is a statement that *at the end* of the knowledge acquisition process, either “A” or “non-A” will result, and no third alternative can be expected. It does not exclude that *in the interim*, a situation of uncertainty may occur. Nagarjuna’s position exploits this confusion in people’s minds.

Nagarjuna interprets the limitation implied by the dilemma “A or non-A” as an arbitrary ‘dualism’ on the part of ordinary thinkers<sup>91</sup>. It only goes to show that he misunderstands formalization (or he pretends to, in an attempt to confuse gullible readers). When logicians use a variable like “B” and allow that “non-A and B” and “A and non-B” are both in principle possible, they do not intend that as a generality applicable to *all* values of B (such as “A”), but only as a generic statement applicable to *any consistent* values of B. In the specific case where B = A,

---

<sup>91</sup> It is misleading to call this a ‘duality’ or ‘dichotomy’, as Buddhists are wont to do, because it suggests that a unitary thing was arbitrarily cut into two – and incidentally, that it might just as well have been cut into four. But, on a perceptual level, there is no choice involved, and no ‘cutting-up’ of anything. A phenomenon appearing is *one single* thing, call it ‘a’ (a proper name, or an indicative ‘this’), and not a disjunction. The issue of ‘dichotomy’ arises only on a conceptual level. *Negation* is a rational act, i.e. we can only speak of ‘non-a’, of what does not appear, by first bringing to mind something ‘a’, which previously appeared (in sensation or imagination). In *initial conceptualization*, two phenomena are compared and contrasted, to each other and to other things, in some respect(s); the issue is then, are they similar enough to each other and different enough from other things to be judged ‘same’ and labeled by a general term (say ‘A’), or should they be judged ‘different’ or is there an uncertainty. At *the later stage of recognition*, we have to decide whether a third phenomenon fits in the class formed for the previous two (i.e. falls under ‘A’) or does not fit in (i.e. falls under ‘non-A’) or remains in doubt. In the latter case, we wonder whether it is ‘A’ or ‘non-A’, and forewarn that it cannot be both or neither.

the said two combinations have to be eliminated because they are illegal (i.e. breach two of the laws of thought).

The above-stated property of symbols, i.e. their applicability only conditionally within the constraints of consistency, is evident throughout the science of formal logic, and it is here totally ignored by Nagarjuna. His motive of course was to verbalize and rationalize the Buddha's doctrine that the ultimate truth is beyond *nama* and *rupa*, name and form (i.e. discrimination and discourse), knowable only by a transcendental consciousness (the Twofold Truth doctrine). More precisely, as Cheng emphasizes, Nagarjuna's intent was to show that logic is inherently inconsistent and thus that reason is confused madness to be rejected. That is, he was (here and throughout) not ultimately trying to defend a tetralemma with B equal to A – or even to affirm that things are both A and non-A, or neither A nor non-A – but wished to get us to look altogether beyond the distinctions of conceptualization and the judgments of logic.

But as above shown he does not succeed in this quest. For his critique depends on a misrepresentation of logical science. He claims to show that logic is confused and self-contradictory, but in truth what he presents as the thesis of logical science is not what it claims for itself but precisely what it explicitly forbids. Furthermore, suppose logical theory did lead to contradictions as he claims, this fact would not lead us to its rejection were there not also a tacit appeal to our preference for the logical in practice. If logic were false, contradictions would be acceptable. Thus, funnily enough, Nagarjuna appeals to our logical habit in his very recommendation to us to ignore logic. In sum, though he gives the illusion that it is reasonable to abandon reason, it is easy to see that his conclusion is foregone and his means are faulty.

## 2. Neither real nor unreal

But Nagarjuna also conceives ultimate reality (“emptiness”<sup>92</sup>) as a “middle way”<sup>93</sup> – so that the world of experience is neither to be regarded as real, nor to be regarded as unreal (“there is nothing, neither mental nor non-mental, which is real” and it “cannot be conceived as unreal,” reports Cheng). In this context, Nagarjuna is clearly relying on one of the above-mentioned logically impossible disjuncts, namely “neither A nor non-A” (be it said in passing). I want to now show why Nagarjuna’s statement seems superficially reasonable and true.

As I have often clarified and explained<sup>94</sup>, knowledge has to be regarded or approached phenomenologically (that is the only consistent epistemological thesis). We have to start by acknowledging and observing *appearances*, as such, without initial judgment as to their reality or illusion. At first sight all appearances seem *real* enough. But after a while, we have to recognize that some appearances conflict with other appearances, and judge such appearances (i.e. one or more of those in conflict) as *illusory*. Since there is nothing in our ‘world’ but appearances, all remaining appearances not judged as illusions (i.e. so long as they are not logically invalidated by conflicts with other appearances) maintain their initial status as realities.

That is, the distinction between appearances as realities or illusions emerges within the world of appearances itself, merely classifying some this way and the rest that way. We have no concept of reality or illusion other than with

---

<sup>92</sup> Beyond consciousness of “Shunyata” is a more vivid awareness called “Mahamudra”, according to Chögyam Trungpa, in *Illusion’s Game*. But such refinements need not concern us here.

<sup>93</sup> See Cheng, pp. 38-39, on this topic. He there refers to MT XIII:9a and XVIII:7.

<sup>94</sup> See my *Future Logic*, ch. 60-62, and later essays on the subject.



reference to appearance. To use the category of reality with reference to something *beyond* appearance is concept stealing, a misuse of the concept, an extrapolation which ignores the concept's actual genesis in the context of appearance. To apply the concept of illusion to *all* appearances, on the basis that some appearances are illusions, is an unjustified generalization ignoring how this concept arises with reference to a specific event (namely, inconsistency between certain appearances and resulting diminishment of their innate credibilities). Moreover, to claim that no appearances are real or that all are illusions is self-defeating, since such claim itself logically falls under the category of appearance.

The illusory exists even though it is not reality – it exists as appearance. The real is also apparent – some of it, at least. Therefore, appearance per se is neither to be understood as reality (since some appearances are illusory), nor can it be equated to illusion (since not all appearances have been or can be found illusory). Appearance is thus the *common ground* of realities and illusions, their common characteristic, the dialectical synthesis of those theses and antitheses. It is a genus, they are mutually exclusive species of it. (The difference between appearance and existence is another issue, I have dealt with elsewhere – briefly put, existence is a genus of appearance and non-appearance, the latter concepts being relative to that of consciousness whereas the former is assumed independent.)

None of these insights allows the conclusion that appearances are “neither real nor unreal” (granting that ‘unreal’ is understood to mean ‘non-real’). All we can say is that some appearances are real and some unreal. Formally, the correct logical relation between the three concepts is as follows. *Deductively*, appearance is implied by reality and illusion, but does not imply them; for reality and illusion are contradictory, so that they cannot both be true and they cannot both be false. Moreover, *inductively*,

appearance implies reality, until and unless it is judged to be illusion (by virtue of some inconsistency being discovered).

More precisely, all appearances are initially classed as real. Any appearance found self-contradictory is (deductively) illusory, and its contradictory is consequently self-evident and (deductively) real. All remaining appearances remain classed as real, so long as uncontested. Those that are contested have to be evaluated dynamically. When one appearance is belied by another, they are both put in doubt by the conflict between them, and so both become initially *problematic*. Thereafter, their relative credibilities have to be tentatively weighed in the overall context of available empirical and rational knowledge – and repeatedly reassessed thereafter, as that context develops and evolves. On this basis, one of these appearances may be judged more credible than the other, so that the former is labeled *probable* (close to real) and the latter relatively *improbable* (close to illusory). In the limit, they may be characterized as respectively effectively (inductively) real or illusory. Thus, reality and illusion are the extremes (respectively, 100% and 0%) in a broad range of probabilities with many intermediate degrees (including problemacy at the mid-point).

To be still more precise, *pure percepts* (i.e. concrete appearances, phenomena) are never illusory. The value-judgment of ‘illusory’ properly concerns concepts (i.e. abstract appearances, ‘universals’) only. When we say of a percept that it was illusory, we just mean that we misinterpreted it. That is, what we initially considered as a pure percept, had in fact *an admixture of concept*, which as it turned out was erroneous. For example, I see certain shapes and colors in the distance and think ‘here comes a girl on a bike’, but as I get closer I realize that all I saw was a pile of rubbish by the roadside. The pure percept is the shapes and colors I see; the false interpretation is ‘girl on

bike’, the truer interpretation is ‘pile of rubbish’. The initial percept has not changed, but my greater proximity has added perceptual details to it. My first impression was correct, only my initial judgment was wrong. I revise the latter concept, not through some superior means to knowledge, but simply by means of *further perception and conception*.

Strictly speaking, then, perception is never at issue; it is our conceptions that we evaluate. It is in practice, admittedly, often very difficult to isolate a percept from its interpretation, i.e. from conceptual appendages to it. Our perception of things is, indeed, to a great extent ‘eidetic’. This fact need not, however, cause us to reject any perception (as many Western philosophers, as well as Buddhists, quickly do), or even all conception. The conceptual ‘impurities’ in percepts are not necessarily wrong. We know them to have been wrong, when we discover a specific cause for complaint – namely, a logical or experiential contradiction. So long as we find no such specific fault with them, they may be considered right. This just means that we have to apply the rules of adduction<sup>95</sup> to our immediate interpretations of individual percepts, just as we do to complex theories relative to masses of percepts. These rules are universal: no judgment is exempt from the requirement of careful scrutiny and reevaluation.

Now, judging by Cheng’s account and certain quotations of Nagarjuna therein, we could interpret the latter as having been trying to say just what I have said. For instance, Cheng writes<sup>96</sup>: “What Nagarjuna wanted to deny is that empirical phenomena... are absolutely real.... However, [this] does not mean that nothing exists. *It does not nullify*

---

<sup>95</sup> Adduction treats all conceptual knowledge as hypothetical, to be tested repeatedly – in competition with all conceivable alternative hypotheses – with reference to all available logic and experience.

<sup>96</sup> P. 42.

*anything in the world*" (my italics). I interpret this non-nullification as an acknowledgment of appearance as the minimum basis of knowledge. Nagarjuna may have had difficulties developing an appropriate terminology (distinguishing existence, appearance and reality, as I do above), influenced no doubt by his penchant for paradoxical statements seeming to express and confirm Buddhist mystical doctrine.

But if that is what he meant, then he has not succeeded to arrive at a "middle way" (a denial of the Law of the Excluded Middle), but only at a "common way" (a granted common ground). As far as I am concerned, that is not a meager achievement – the philosophical discovery of phenomenology! But for him that would be trivial, if not counterproductive – for what he seeks is to deny ordinary consciousness and its inhibiting rationales, and to thereby leap into a different, higher consciousness capable of reaching transcendental truth or ultimate reality.

It is interesting to note that the Madhyamika school's effective denial of reality to all appearance was not accepted by a later school of Mahayana philosophy, the Yogacara (7<sup>th</sup>-8<sup>th</sup> cent. CE). Cheng describes the latter's position as follows<sup>97</sup>: "Every object, both mental and non-mental, may be logically or dialectically proven illusory. But in order to be illusory, there must be a certain thought that suffers from illusion. *The very fact of illusion itself proves the existence and reality of a certain consciousness or mind.* To say that everything mental and non-mental is unreal is intellectually suicidal. The reality of something should at least be admitted in order to make sense of talking about illusion" (italics mine). That is the tenor of the phenomenological argument I present above, although my final conclusion is clearly not like Yogacara's, that everything is consciousness or mind (a type of Idealism),

---

<sup>97</sup>

P. 25.

but leaves open the possibility of judging and classifying appearances as matter or mind with reference to various considerations.

The Madhyamika rejection of ‘dualism’ goes so far as to imply that “emptiness” is not to be found in nirvana, the antithesis of samsara (according to the earlier Buddhist viewpoint), but in ‘neither samsara nor nirvana’. In truth, similar statements may be found in the Pali Canon, i.e. in the much earlier Theravada schools, so that it is not a distinctly Mahayana construct. The difference is one of emphasis, such statements, relatively rare in the earlier period, are the norm and frequently repeated in the later period. An example may be found in the *Dhammapada*, a sutra dating from the 3<sup>rd</sup> cent. BCE<sup>98</sup>, i.e. four or five hundred years before Nagarjuna. Here, samsara is likened to a stream or this shore of it, and nirvana to the further shore; and we are told to get beyond the two.

*When you have crossed the stream of Samsara, you will reach Nirvana... He has reached the other shore, then he attains the supreme vision and all his fetters are broken. He for whom there is neither this nor the further shore, nor both....*

Such a formula is legitimate if taken as a warning that *pursuing* nirvana (enlightenment and liberation) is an obstacle to achieving it, just a subtle form of samsara

---

<sup>98</sup> This is supposedly the date of composition, though the translator, Juan Mascaro, in his Introduction, states “compiled” at that time, thus seeming to imply an earlier composition. It is not clear in that commentary when the sutra is estimated to have been first written down. And if it was much later, say in the period of crystallization of Mahayana thought, say in 100 BCE to 100 CE, the latter may have influenced the monks who did the writing down. See ch. 26 (383-5) for the quotation.

(ignorance and attachment); there is no contradiction in saying that *the thought of nirvana* as a goal of action keeps us in samsara – this is an ordinary causal statement. The formula is also logically acceptable if taken as a reminder that no word or concept – not even ‘samsara’ or ‘nirvana’ – can capture or transmit the full meanings intended (i.e. ‘not’ here should more precisely be stated as ‘not quite’). There is also no contradiction in saying that one who has attained nirvana does not need to leave the world of those locked in samsara, but can continue to exist and act in it though distinctively in a way free of attachment.

But it would be a contradiction in terms to speak of ‘emptiness’ as ‘neither samsara nor nirvana’, given that nirvana as a concept is originally defined as non-samsara; the truth cannot be a third alternative. At best, one could say that emptiness is a higher level of nirvana (in an enlarged sense), which is not to be confused with the lower level intended by the original term nirvana, nor of course with samsara. In that case, nirvana (in a generic sense of the term, meaning literally non-samsara) includes both a higher species and a lower one; and the statement ‘neither samsara nor lower-nirvana’ is then compatible with the statement ‘higher nirvana’. There is a big difference between rough, poetic, dramatic language, and literal interpretation thereof.

## 15. More Buddhist antinomic discourse

*From Ruminations 5, and Logical and Spiritual Reflections 3.10.*

### 1. The ‘I have no thesis’ thesis

a. The Buddhist<sup>99</sup> philosopher Nagarjuna (India, c. 150-250 CE) attacked every thesis he regarded as rational by every means he regarded as logical, and declared his own discourse immune from scrutiny and criticism, by saying (according to one translation):

*“If I had a thesis, I would be at fault; since I alone have no thesis, I alone am without fault” (VV 29)<sup>100</sup>.*

---

<sup>99</sup> Needless to say, the following comments are not an attack on Buddhism, but on the rhetoric of Nagarjuna. Buddhism is not well served by such games. I think of Nagarjuna whenever I read v. 306 of the *Dhammapada*: “He who says what is not... and he who says he has not done what he knows well he has done... sinned against truth”. For me, he is just a philosopher like any other; his interest in Buddhism is incidental (as is his saintly status in the eyes of many).

<sup>100</sup> Nagarjuna in *Vigraha Vyavartani (Averting the Arguments)*, verse 29. The translation used here is given by ‘Namdrol’ in the E-Sangha Buddhism Forum (<http://www.lioncity.net/buddhism/index.php?s=d8946a5bcb1f56f3e9e21a108125823f&showtopic=5604&st=100&#entry82577>). Note however that the word “alone” in this translation may not be in the original, judging by other translations I have seen, even though it does seem to be Nagarjuna’s intent.

The first aspect of Nagarjuna's statement is a brazen **claim to have no claim**. This is of course self-contradictory. Every proposition that claims to be meaningful and true (whether about some experience or about abstraction, whether positive or negative) is an assertion, a claim. To pretend making no claim even as one plainly makes one is a breach of the law of identity: it is denying that a fact is a fact.

There is no logical way to deny or criticize the theses or methodologies of others without opening one's own discourse to evaluation. All denial or criticism is discourse, and all discourse is subject to logical review. To pretend the logical possibility of dispensation is dishonest (and if such pretense implicitly is bad enough, it is all the more dishonest if made explicitly).

Nagarjuna's discourse was, in fact (as I show in *Buddhist Illogic*), shock full of fallacious arguments, a mere parody of logic posing as logic. But he knew that people untrained in logic would fall for it, and he sealed their intellectual fate with the said eyewash claim. To neutralize further discussion, he misled them into believing he had simply shown up the logical absurdity of logic, and all doctrines based on it, but had himself posited no methodology or doctrine of his own.

Not only was his alleged refutation of reason full of errors of reasoning, but his concluding 'no-claim claim' was also a mockery of logic and sincerity. He, of course, just says 'I make no claim' – and he persistently denies that this statement constitutes a claim. I call that shameless psychological manipulation, motivated by one-upmanship. He cynically takes advantage of the credulity of some people, to dominate them intellectually.

The second aspect of Nagarjuna's above statement can be viewed as a 'soft' version of the liar paradox, since he tells



us: **everyone but me is in error**. Although such a statement is not in itself inconsistent (God could conceivably utter it truthfully) – it is logically open to doubt due to being *self-exempting*.

Effectively, it says: ‘I am the only human who has knowledge; I know everyone else is incapable of true knowledge’. Only a fool is tricked by such an unsubstantiated claim to privilege. Reason regards all people as technically within range of knowledge given enough effort, even if they do not all fulfill their potential equally. Reason demands that discourse be reasoned and fair – i.e. based on *common general norms* as to how truth and falsehood are to be determined.

If Nagarjuna were basing his criticism of ordinary human means to knowledge on a claim to have attained a ‘*higher level*’ of *consciousness* (i.e. Buddhist enlightenment or Biblical prophecy), we could not convincingly oppose him (being unable to prove or disprove such experiential claims). But he is not using such a claim as his basis – he is attempting to debunk reason through *ordinary* logical discourse. In that case, he is fair game for logic.

The statement of infallibility is then seen as manifest arrogance, a lack of respect for other thinkers. By saying ‘I alone am exempt from any criticism’ the author aggressively grants himself a special dispensation: he alone is endowed with the way to knowledge; everyone else is an idiot or a dishonest person. It is totalitarian, dictatorial speech.

Compare this dismissive ‘you all know nothing’, to the self-inclusive statement ‘I (or we) know nothing’. The latter – even though it implies ‘I know that I know nothing’ and is therefore self-inconsistent – is at least modest; so much so, that such admission is widely considered a mark of wisdom (and it is commendable, in modified form, i.e. as ‘I know *close to* nothing, very little’).

Self-exemption is a hidden form of self-inconsistency, because it resorts to *a double standard*. The one making such a claim presents superficially rational arguments against human experience and logic, but does not ask himself or tell us how he (an ordinary human) managed (using the very cognitive means he rejects) to attain such allegedly true knowledge. The author criticizes others, but does not equally well criticize himself.

This is a fallacious mode of thought often found among would-be skeptical philosophers. It comes in many subtle forms. It is wise to always be on the lookout for such practices, applying the reflexive test here demonstrated.

b. Looking at Nagarjuna's above statement in more detail, the following may be added.

To begin with, what is meant here by "*having a thesis*"? This refers to any explicit or even wordless belief, any clear or even vague opinion upheld (considered to constitute knowledge), any proposition one advocates or implicitly logically condones. The subject that Nagarjuna is here discussing is any outcome of human rational cognition, any belief, opinion or doctrine that one may arrive at, rightly or wrongly, by means of ordinary consciousness, i.e. through experience, negation, abstraction, hypothesizing, inductive or deductive argument.

And what is meant here by "*being at fault*"? This refers to making a mistake in the course of observation or reasoning, so that some thesis one has adhered to is in fact an illusion rather than a reality, false rather than true, erroneous instead of correct.

How do we know the status appropriate to a thesis? We know it (I suggest) *by holistic application of the whole science of logic to the totality of the data of experience*. Our concepts of cognitive right or wrong are themselves all constructed by logic and experience, without appeal to some extraordinary outside justification (like prophetic

revelation or mystical realization, or simply the authority of some great personage or of a religious document or institution).

Now, Nagarjuna is evidently well aware of all that, but is intent on annulling the independent reliability of ordinary experience and reason. His strategy and tactics to this end, in all his discourse, as I have shown throughout my *Buddhist Illogic*, is to give the impression (however paradoxical) that logic may be invalidated by means of logic. And this twofold sentence of his, “If I had a thesis, I would be at fault; since I alone have no thesis, I alone am without fault”, fits neatly into his destructive philosophical programme.

On the surface, this sentence might be construed as a single argument:

If X (a proposition is proposed), then Y (an error is made)

but not X (no proposition)

therefore, not Y (no error)

Although the above apodosis is logically invalid, since it denies the antecedent to deny the consequent, Nagarjuna is not above letting it pass without comment, knowing it will suffice to convince some people, although he is well aware that the logically trained will spot it and object. But for the latter audience, he reserves a subtler form of manipulation.

It has to be seen that the purpose of this famous Verse 29 in Nagarjuna’s discourse is designed *to make a show of logical consistency*. He wants by means of it to give the impression that his anti-rational discourse is justifiable, that it has the stamp of approval of logic. Yes, he is actually attacking logic; but at the same time, he has to pretend to use it, because he knows this measure is required to

convince people. For most people, a veneer of logic (i.e. mere rhetoric) suffices to put their reason's critical faculty at rest. We shall now see how he goes about this task.

The first part of Nagarjuna's statement, viz. "If I had a thesis, I would be at fault", is not intended (as some have assumed) as a justification for his overall discourse. It is not placed here in his discourse as an argument with intrinsic force, which directly buttresses or proves his philosophy. It is certainly not an obvious logical principle, or axiom, which everyone would agree on without objection, from which his discourse can be inferred or even generalized. No – it is itself an inference and application from Nagarjuna's main thesis, namely the claim that 'All human knowledge based on ordinary experience and reason is necessarily erroneous'.

The latter underlying claim is his major premise in a (here tacit) productive eduction, i.e. one that deduces a particular hypothetical proposition from a more general categorical one<sup>101</sup>. This argument is formally valid, running as follows:

All X (opinions) are necessarily Y (erroneous);

therefore,

If this is X (a proposition is proposed), then this is Y (an error is made).

In this way, the first part of Nagarjuna's statement is made to seem something inferred, rather than an arbitrary claim. It is cunningly presented as an application of already admitted information, rather than as an isolated assertion. Granting the premise, the conclusion indeed logically follows (this is the veneer of logic) – but has the premise already been granted? No. Also note, once the conclusion

---

<sup>101</sup> See *Future Logic*, chapter 29.3.

is seemingly drawn, it can by generalization be used to reinforce the premise; although this is a circularity, it works psychologically.

Moreover, Nagarjuna manages through this implicit productive argument to pretend he is being consistent with himself: he is telling us, effectively: ‘See, I am not just attacking other people’s knowledge, but am prepared to apply the same stringent critique to my own!’ This virtuous declaration is of course dust in your eyes, because he is not here putting the broader principle in doubt but merely reaffirming it. He has nowhere established that ‘All propositions are false’. His is a pseudo-logical posture.

As the next part of his statement clarifies, he does not consider his discourse as falling under the critical rule he has formulated. The proposition “If I had a thesis, I would be at fault” is a counterfactual hypothetical; his own discourse is never made into an issue open to debate. It seems open-minded, but it is a foregone judgment. His intention is to ‘avert all arguments’ and place himself at the outset outside the fray. He seemingly at first admits and then vehemently denies that his own discourse is a product of ordinary consciousness. This convoluted avoidance of cognitive responsibility has fooled many a poor soul.

Moving on, now, to the second part of Nagarjuna’s statement, viz. “since I alone have no thesis, I alone am without fault”. As already pointed out, this can be viewed as the minor premise and conclusion of an invalid apodosis in which the first part of the statement is the major premise. But we could also more generously assume that Nagarjuna intended a valid apodosis, using as its tacit major premise the obvious proposition: ‘If one has no thesis, one cannot make a mistake’.

It can be correctly argued that this premise was left tacit simply because it is so obvious to and readily granted by everyone. It is indeed true that if one ventures no utterance,

thought or even intention, if one holds no opinion, makes no claim to knowledge, if one remains inwardly and outwardly silent, one will never make any errors. *For the status of truth or falsehood is only applicable to meaningful assertions.*

A stone is never in error, because it has no thoughts. Likewise, a thoughtless person may by his or her ignorance, blindness or stupidity make many errors of living, but makes no error in the logical sense of having proposed an inappropriate proposition. All that is so obvious (and vacuous) no debating it is necessary. The following apodosis is thus implicit in Nagarjuna's declaration:

If not X (no proposition is proposed), then Y (no error is made)

but not X (no proposition)

therefore, not Y (no error)

This argument has a true major premise, as well as a valid form. This gives his discourse a veneer of logic again, helping him to persuade more victims. However, his minor premise remains well open to doubt, and decisively deniable! (As a consequence of which, his conclusion is of course also open to doubt.) He takes it for granted that he 'has no thesis' – but this claim is far from granted already. The tacit major premise acts as a smokescreen for the minor premise.

Moreover, note, although 'being correct' implies 'not being at fault', the reverse is not necessary. Nagarjuna suggests that his alleged faultlessness implies the correctness of his position, but it does not follow! Only if his criticism of all opposing theses was correct (which is by no stretch of the imagination true), and his thesis was

not liable to similar criticism and was therefore the only leftover logical possibility, would such inference be drawn. Nagarjuna does indeed 'have a thesis'. His main thesis, the goal of his whole philosophical discourse, is as already mentioned the claim that 'All human knowledge based on ordinary experience and reason is necessarily erroneous'. This, for a start, qualifies as a thesis – boy, it is a big skeptical thesis, full of negative implications. It is a principle of logic that *to deny any thesis is to affirm an opposite thesis*. His claim that his doctrine is not a thesis, in the minor premise here, is mere arbitrary assertion.

Furthermore, he 'has a thesis' every time he makes a specific assertion of any kind, including the assertion under scrutiny here, viz. "If I had a thesis, I would be at fault; since I alone have no thesis, I alone am without fault". Note that Nagarjuna thinks that making a negative statement is somehow 'not having a thesis' – but the polarity of a statement does not diminish the need for justification; if anything, one can argue that on the contrary negative statements are harder to establish than positive ones!

And we should strictly include as 'theses' of his not only such explicit statements, but also all the implicit assumptions and suggestions within his discourse (like the implicit major premise and resulting apodosis we have just highlighted). It makes no difference whether these explicit, or unstated and unadmitted, items constitute information or logical method, content or process.

For all these elements of discourse, be they spoken or otherwise intended, in all fairness fit in our common understanding and definition as to what it means to 'have a thesis'. For none of these categorical or hypothetical propositions (except perhaps 'if silence, no error') is self-evident. They did not arise *ex nihilo* in Nagarjuna's mind, ready-made and self-justified.

They are all complex products of ordinary human cognition, based on experience and produced by reason (even if, in Nagarjuna's case, the mind involved is deranged). They undeniably together form a specific philosophy, a theory of logic, an epistemology and ontology. The mere fact that we can (as here done) at all consider and debate them is proof that they are 'theses'.

The law of identity (A is A) must be maintained: facts are facts and it is no use pretending otherwise. Nagarjuna may eternally refuse the predicate of "having a thesis", but we confidently insist on it. His arguments have in no way succeeded in averting this just and true judgment. Consequently, his doctrine is self-contradictory. Not only does he 'have a thesis', but since his thesis is that 'to have a thesis is to be in error', he has (by its own terms) to be recognized as being in error.

Thus, to end it: Nagarjuna's statement "If I had a thesis, I would be at fault; since I alone have no thesis, I alone am without fault" weaves a complicated web of deception. It misleads, by means of subtle ambiguities and superficial imitations of logic. Once its dishonesty is revealed, it should be decidedly rejected.

The mere historic fact that Nagarjuna is famous and admired by many does not justify hanging on to his doctrine *ad nauseam*, trying ex post facto to find ways to make it consistent with logic. Celebrity is not proof of some hidden truth – it is vanity. Most who do so are merely grasping for reflected glory. Anyway, attachment to authority is argument *ad hominem*. The religious and academic 'groupies' who gave him and perpetuate his authority are not logically competent, however numerous they be. It is a case of the blind leading the blind.

c. Nagarjuna defends his 'non-thesis' idea in the next verse (VV 30), describing it as "*a non-apprehension of*



*non-things*” (according to one translation<sup>102</sup>). Now, this is a very funny phrase. To the impressionable, it sounds very deep, pregnant with meaning. It seems to suggest this man has some privileged higher way of knowledge that goes beyond ordinary experience and reasoning.

But in truth, taken literally, we are all quite capable of “non-apprehension of non-things” and daily practice it, for the simple reason that non-things cannot be apprehended! Logically, this is all this phrase means, note well. What then is the old fox up to, here?

Nagarjuna is trying to project his ‘not having a thesis’ position as far as logically possible from our plebian ‘having a thesis’ – i.e. from ordinary consciousness, which consists in ‘the apprehension of things’. He has logically only three alternatives to choose from:

- the ‘non-apprehension of things’ (unconsciousness);
- the ‘apprehension of non-things’ (an otherworldly consciousness);
- or the ‘non-apprehension of non-things’.

Having a marked taste for one-upmanship and dramatic extremes, Nagarjuna of course chose the third of these terms as his vehicle. Even though the obvious sense of this phrase is puerile, it has poetic breadth and appeal. It seems

---

<sup>102</sup> By Frederick J. Streng. The full text of his translation seems to be that posted in the Internet at: <http://www.orientalia.org/article491.html>. Note that the phrase “non-apprehension of non-things” is considered an incorrect translation by Plamen Gradinarov. However, while willing to admit the latter’s objection, I do not agree that Streng’s freer translation is entirely inadmissible. In my view, it may not be literally precise, but it captures Nagarjuna’s paradoxical spirit and intent. See our discussion of this issue at <http://nyaya.darsana.org/topic3.html>. In any case, even if the phrase “non-apprehension of non-things” is best not relied on, the criticisms of Nagarjuna in the present section can still be proposed on other grounds.

to imply ‘knowledge without consciousness’ and ‘consciousness of the unknowable’ all at once.

Thus, his ‘non-apprehension’ is a mix of apprehension and non-apprehension, or something else again. And likewise, his ‘non-things’ are things of some sort as well as non-things, or perhaps something quite other still.

In other words, the negative terms in the phrase “non-apprehension of non-things” are not intended by Nagarjuna nor received by his disciples and students as mere negations of the corresponding positive terms, but as **paradoxical terms**, which may (in accord with the tetralemma schema) be all at once *positive or negative or both or neither*.

It is (and isn’t) ‘apprehension/non-apprehension of things/non-things’ all in one.

Nagarjuna stands out in the history of world philosophy as the most unabashed opponent of the laws of thought. Not only does he freely use self-contradictory or middle-including propositions, but he even makes use of terms loaded with contradiction and inclusion of a middle.

Now, some people might say: ‘what is wrong with that?’ They will argue: ‘the real world is extremely subtle, and we can only ever hope to express it in thought very approximately; Nagarjuna is only trying to take this uncertainty into consideration within his discourse; the laws of thought are just arbitrary demands, making us force our thoughts into prejudicial straightjackets’.

But logical laxity is not the proper attitude in the face of an extremely complex and hard to express real world. It is precisely because of the great difficulty of the cognitive task at hand that one is called upon to be very clear and careful. Avoiding checks and balances on our judgments does not increase their efficiency but makes them less reliable.

In the case under consideration, if Nagarjuna does indeed have some privileged form of otherworldly consciousness, he can just say so. The laws of thought in no way forbid him to posit such a claim. He does not need to beat about the bush, and pretend to have something unspeakable and not subject to peer review. He can and should be forthright, and defend his position in an equitable way like everyone else.

If he considers the terms ‘apprehension’ and ‘things’ to have some intrinsic logical flaw, he can argue his case openly; he does not need to engage in allusion, suggestion and fallacious argument. Most of us thinkers are open-minded and willing to correct our errors: if these terms are flawed, we are not attached to them; we are flexible, ready to modify or replace them as logically necessary in the light of new evidence and reasoning.

But Nagarjuna is like an accused, who when forced to appear in court refuses to admit his identity, or recognize the authority of the law and the judges, or plead guilty or not guilty, or argue the defense of his case. Worse still, in utter contempt of the court, he does not even admit his refusal to be a refusal – he calls it a ‘non-thesis’. Does that stop court proceedings or make the court declare him innocent? Surely not.

Nagarjuna misunderstands the nature of negation. He thinks that if one person says ‘X’ and another says ‘not X’, the onus of proof is on the first more than on the second. He considers that making a positive statement is more logically demanding than making a negative one. *He imagines in his confusion that saying ‘no’ is equivalent to saying nothing, i.e. to not saying anything.* Most logicians would disagree with him, and argue that any thesis put forward (even if only by insinuation) is equally in need of proof, whatever its polarity.

I would go further and say that, on the contrary, a negative statement is more demanding than a positive one. You can prove a positive statement easily enough, if you point to sufficient evidence in its favor. But how do you prove a negative statement? It is much more difficult, since negatives are not directly experienced but are only experienced by way of the absence of positives. A negative can ultimately only be proved indirectly, by inability to prove any contrary positive.

Thus, in fact, not only does Nagarjuna's alleged self-limitation to negatives not exempt him from proofs, but on the contrary it increases the logical burden upon him. He is right in considering negatives as significantly different from positives, but he does not realize that the difference is to his disadvantage. He claims to have no epistemological or ontological basis, and yet to be able to reject offhand all theories of knowledge and reality. Such a grandiose fanciful claim surely requires much more justification than any other!

It should be stressed, incidentally, that Nagarjuna's "non-apprehension of non-things" should not be interpreted (as some do) as a defense of non-verbal meditative experience or insight. That is not the thrust of his anti-rational philosophy, although its avowed Buddhist affiliation may lead one to suppose so.

If Nagarjuna were a man deeply absorbed in meditation, he would not be writing philosophy. If his intent were to promote meditation, he would simply teach methods of meditation and not stir up verbal disputes. No – this man has philosophical ambitions. Allegedly, these are meant to put into words some of the 'reasoning' that he considered the Buddha to have gone through before attaining enlightenment. Nagarjuna assumes from the start that this 'reasoning' is necessarily anti-logical, a rejection of reason.

But we must see that this assumption is just a prejudice of his distorted mind. He was a philosophical revolutionary – one who believed that reason has to be overturned, to be transcended. But it is more credible to be evolutionary – and to consider meditation as a way for us to keep moving, beyond the limits of discursive thought, without need to deny such thought within its applicable bounds.

To advocate respect for logic is not to foment endless babble, but rather to require that any thought arising be subjected to responsible cognitive evaluation. Logic is possible entirely without words, by means of silent intentions. Even in deep meditation, some sort of ‘reality check’ by means of logic occurs, and this need not involve any words. It is only by this means, no doubt, that a Buddha-to-be may steer himself well clear of common illusions and insane imaginings, towards to full realization.

Contrary to Nagarjuna’s belief, rationality and spirituality are not necessarily in conflict. Reason and meditation are potentially, to some extent, mutually beneficial. It is not thought as such, much less logic, but only excess of thought, particularly irrelevant chatter, which hinders meditative concentration and contemplation. A certain amount of appropriate thinking is often needed to initially position one’s mind for meditation.

d. In fact, as I will now show, the sentence “If I had a thesis, I would be at fault”<sup>103</sup> is a **formal impossibility**. I earlier interpreted and symbolized it as “If X (a proposition is proposed), then Y (an error is made)”, giving the antecedent and consequent two separate symbols, X and Y. But now let us consider these constituents more closely.

---

<sup>103</sup> Two other translations of this sentence confirm and amplify this reading. “If I would make any proposition whatever, then by that I would have a logical error” (Streng). “Should I have put forward any thesis, then the logical defect would have been mine” (Gradinarov).

What does “making an error” mean here? It is not an ordinary predicate. The consequent Y does not merely refer to some error in general, but specifically to an error *in the antecedent* X. Y tells us that X *is wrong*. Therefore, Y formally implies the negation of X, i.e. notX! Granting this, Nagarjuna’s sentence now reads: “**If X, then not X**”, i.e. “If X is true, then X is false” – a paradoxical hypothetical proposition, whose conclusion would be the categorical “X is false” (as earlier suggested).

However, that is not the end of the matter. If we now consider the meaning of X – viz. “a proposition is proposed” – we may fairly suppose it refers to *just any proposition whatsoever*. In that case, the proposition concerned might even be the negation of X; so that we may substitute notX for X throughout the hypothesis. So doing, we obtain “If notX, then not notX”, i.e. “**If not X, then X**”, or in other words “If X is false, then X is true”. This is also, of course, a paradoxical proposition, whose formal conclusion is “X is true”.

We thus – by means of a universal reading of “having a thesis”, as inclusive of “not having a thesis” – now have, not only a single paradox, but a *double paradox*! That is, our conclusion is not only that X is false, but that X is both true and false. The latter conclusion is of course contrary to the law of non-contradiction, as in the case of the liar paradox.

This means that Nagarjuna’s statement is a formal impossibility: it is a contradiction in terms; it is not only false, but meaningless. It does not constitute legitimate discourse at all, let alone a tenable philosophical position or theory. The words or symbols used in it are logically not even conceivable, so it is as if he is saying nothing. He seems to be saying something intelligible, but it is an illusion.

Now, it may be objected that Y does not necessarily mean that X *is* wrong, but could merely mean that X *could be* wrong. That is, “making an error” could be taken to mean that X is uncertain rather than definitely refuted. In that case, we would have the following two hypotheses: “If X, possibly not X” and “If not X, possibly X”; or in one sentence: “Whether X or not X is proposed, the outcome is uncertain”. Indeed, this more modal, ambiguous posture may well be considered as Nagarjuna’s exact intent (which some have interpreted as noncommittal ‘illocution’).

At first sight, due to the use of vague words or of symbols, this objection may seem credible and the contradictory conclusions involved apparently dissolved. But upon reflection, there is still an underlying conflict: to affirm X, or to deny it, is contrary to a position that neither affirms nor denies X. An assertoric statement (affirming or denying X) is *incompatible with* a problematic statement (saying X may or may not be true). One cannot at once claim to *have* knowledge (of X, or of not X) and claim to *lack* it (considering the truth or falsehood issue open). This is as much a contradiction as claiming the same thing (X) true and false.

Someone unacquainted with the logic of hypothetical propositions might now object that X, or notX, is only proposed hypothetically in the antecedent, and so may well be problematic in the consequent. But this is a logically untenable objection, due to the process of addition (described in the chapter on formal logic); i.e. due to the fact that “If X, then Y” implies “If X, then (X and Y)”. In the present case, this means: “If X is asserted, then X is both asserted and uncertain”. It suffices for the contradiction to occur conditionally, as here, for the condition to be disproved; therefore, our conclusion *is quite formal*: “X cannot be asserted”. QED.

Someone could here, finally, object that the certainty in the antecedent and the uncertainty in the consequent may not be simultaneous, and so not produce a logical conflict. Such objection would be valid, granting that a thought process separated the beginning and end of the hypothetical proposition. However, in the case under scrutiny, Nagarjuna is clearly stating that in the very act of “proposing something”, one would be “making an error”; i.e. the error is nothing other than the proposing, itself. So, no time separation can credibly be argued, and Nagarjuna’s thesis remains illogical.

Note that all the present discussion has concerned only the first part of verse 29, i.e. the major premise “If I had a thesis, I would be at fault”. We have found this hypothetical proposition logically faulty, irrespective of whether Nagarjuna admits or refuses to acknowledge that he “has a thesis”. So, let us now reconsider this minor premise of his, and his conclusion that he “is not at fault”.

We have here introduced a new twist in the analysis, when we realized that “If X, then Y” (understood as “If X, then not X”) implies “If not X, then Y” (since the latter is implied by “If not X, then X”, which is implied by the former by replacing X with notX). So, now we have a new major premise for Nagarjuna, namely “If not X, then Y”, meaning: “If I do not have a thesis, I will be at fault”.

Taking this implied major premise with Nagarjuna’s own minor premise, viz. “I have no thesis” – the conclusion is “I am at fault”. This conclusion is, note, the opposite of his (“I am not at fault”). Thus, even though Nagarjuna boasts his thinking is faultless, it is demonstrably faulty!

For – simply put, leaving aside all his rhetoric – all he is saying is: “no thesis is true”; it is just another version of the liar paradox. And his attempt to mitigate his statement, with the afterthought “except my thesis”, is logically merely an additional statement: a particular case that falls



squarely under the general rule. Moreover, before an exception can be applied, the rule itself must be capable of consistent formulation – and this one clearly (as just shown) is not.

Note lastly, none of this refutation implies that silence is impossible or without value. If (as some commentators contend) Nagarjuna's purpose was to promote cessation of discourse, he sure went about it the wrong way. He did not need to develop a controversial, anti-logical philosophy. It would have been enough for him to posit, as a psychological fact, that (inner and outer) silence is expedient for deep meditation.

## 2. Calling what is not a spade a spade

Buddhism, no doubt since its inception, has a mix of logic and illogic in its discourse. Looking at its four main philosophical schools, Abhidharma, Prajnaparamita, Madhyamika and Yogacara, the most prone to discard the three laws of thought (i.e. Identity, Non-contradiction, Exclusion of the middle) was Madhyamika<sup>104</sup>. But this trend was started in the earlier Prajnaparamita, as examples from the *Diamond Sutra*<sup>105</sup> show.

---

<sup>104</sup> See my work *Buddhist Illogic* on this topic, as well as comments on Nagarjuna's discourse in my *Ruminations*, Part I, chapter 5. I must stress that my concern, throughout those previous and the present critiques, is not to reject Buddhism as such, but to show that it can be harmonized with reason. I consider quite unnecessary and counterproductive, the attitude of many Buddhist philosophers, who seemingly consider Realization (i.e. enlightenment, liberation, wisdom) impossible without rejection of logic. My guiding principle throughout is that they are quite compatible, and indeed that reason is an essential means (together with morality and meditation) to that desirable end.

<sup>105</sup> Judging by its Sanskrit language, the centrality of the bodhisattva ideal and other emphases in it, this sutra is a Mahayana text. It is thought to have been composed and written in India about

We do, in this sutra, find samples of valid logical argument. For example, there is a well formed a fortiori argument in Section 12<sup>106</sup>: “wherever this sutra or even four lines of it are preached, that place will be respected by all beings... How much more [worthy of respect] the person who can memorize and recite this sutra...!” But we do also find plain antinomies, like “the Dharma... is neither graspable nor elusive” (said even though not graspable means elusive, and not elusive means graspable).

But the *Diamond Sutra* repeatedly uses a form of argument that, as a logician, I would class as a further twist in the panoply of Buddhist illogic. This states: “**What is called X is not in fact X; therefore, it is called X**” (or sometimes: “**What is called X is truly not X; such is merely a name, which is why it is called X**”).

There are over twenty samples of this argument in the said sutra. Here is one: “What the Tathagata has called the Prajnaparamita, the highest, transcendental wisdom, is not, in fact, the Prajnaparamita and therefore it is called Prajnaparamita.” Here is another: “... what are called beings are truly no beings. Such is merely a name. That is why the Tathagata has spoken of them as beings.”<sup>107</sup>

What I am questioning or contesting here regarding this sort of discourse is only the “*therefore*” or “*which is why*”

---

350 C.E., though at least one authority suggests a date perhaps as early as 150 C.E. For comparison, Nagarjuna, the founder of Madhyamika philosophy, was active circa 150-200 C.E.; thus, this Prajnaparamita text was written during about the same period, if not much later.

<sup>106</sup> Mu Soeng, p. 111.

<sup>107</sup> In Mu Soeng: pp. 145 and 151, respectively. I spotted a similar argument in another Mahayana text: “And it is because for them [the bodhisattvas] training consists in not-training that they are said to be training” (my translation from a French translation) – found in chapter 2, v. 33 of the “Sutra of the words of the Buddha on the Supreme Wisdom” (see Eracle, p. 61).

conjunction<sup>108</sup>. I am not denying that one might call something by an inappropriate name, or even that words can never more than approximate what one really wants to say. But to say that one is naming something X *because* it is not X – this is surely absurd and untenable.

This is not merely ‘not calling a spade a spade’ – it is calling something a spade even while believing it not to be a spade! This is, at least on the surface, contrary to logic. If the label is not applicable, why apply it? Moreover, why boast about this unconscionable inversion, saying “therefore”?

To say that something “is not in fact or truly X” is to imply that the word X has a sense that the thing under consideration does not fit into; in such case, why call that very thing ‘X’ against all logic? Why not just call it ‘not X’ (or coin for it some other, more specific name) and avoid paradox!

Discourse like “such is merely a name” is self-defeating anyway, since in fact it uses names that do convey some meaning. The sentence suggests no words have any valid reference, yet relies on the effectiveness of the words it utilizes to communicate its various intentions. It is a statement that tries to exempt itself from the criticisms it levels at *all* statements as such.

In the examples given above, the argument depends on our understanding of words like ‘Prajnaparamita’ (i.e. perfection of wisdom) or ‘beings’ – and yet at the same time tries to invalidate any such understanding. It cannot therefore be said to communicate anything intelligible.

---

<sup>108</sup> Assuming the translation in this edition is correct, of course (and it seems quite respectable; see p. ix of the Preface). My point is that no logician has ever formally validated such an argument; and in fact it is formally invalid, since the conclusion effectively contradicts a premise.

Without doubt, we cannot adequately express ultimate reality (or God) in words. But it remains true that we can verbally express the fact of ineffability (as just done in the preceding sentence). There is no need to devalue words as such to admit that they have their limits.

Moreover, it is very doubtful that such paradoxical statements (like “name this *X* *because* it is not *X*”) are psychologically expedient to attain enlightenment; they just cognitively confuse and incapacitate the rational mind. Rather than silence the inquiring mind, all they actually do is excite it with subconsciously unanswered questions. Such nonsensical statements are products of an unfortunate fashion that developed in Buddhism at a certain epoch<sup>109</sup>.

That sort of intellectual perversity came to seem profound, as it does to some postmodern thinkers in the West today, precisely because a logical antinomy implies nothing – and that emptiness of meaning is (wrongly) equated with the Emptiness underlying all phenomena. The gaping hole in knowledge left by antinomy gives the illusion of being pregnant with meaning, whereas in fact it is just evidence of ignorance. Note this well.

It should be added that there is indeed a sort of structural paradox in the meditative act – but the *Diamond Sutra*'s habit of ‘calling not a spade a spade’ is not it. The paradox involved is that if we pursue enlightenment through meditation, we cannot hope to attain it, for then our ego (grasping at this transcendental value as at a worldly object) is sustained; yet, meditation is the best way to enlightenment. So, we must ‘just do it’ – just sit and let our native enlightenment (our ‘Buddha nature’) shine forth eventually.

---

<sup>109</sup> Although not entirely absent in the earlier Abhidharma literature and the later Yogacara literature, they are not uncommon in some Prajnaparamita literature (including the *Diamond Sutra*) and rather common in Madhyamika literature.

It should also be reminded that Buddhism is originally motivated by strong realism. It is essentially a striving towards Reality. In this perspective, the Buddhist notion of “suchness” may be considered as a commitment to the Law of Identity. The enlightened man is one who perceives things, in particular and in general, *such as they really are*. This is brought out, for instance, in the following Zen exchange. A monk asked Li-shan: “What is the reason [of Bodhidharma’s coming from the West, i.e. from India to China]”, to which the Zen master replied “Just because things are such as they are”, and in D. T. Suzuki’s commentary that this refers to “Suchness”<sup>110</sup>.

---

<sup>110</sup> *The Zen Doctrine of No-mind*, p. 93.

## Main References

- Cheng, Hsueh-li. *Empty Logic: Madhyamika Buddhism from Chinese Sources*. Delhi: Motilal Banarsidass, 1991.
- Dawkins, Richard. *The Greatest Show on Earth: The Evidence for Evolution*. New York: Free Press, 2009.
- Dhammapada, The*. Juan Mascaro, trans. Harmondsworth, Mx.: Penguin, 1973.
- Eracle, Jean. *Enseignements du Bouddha*. Paris: Librio, 2005.
- Kneale, William and Martha. *The Development of Logic*. Oxford, London: Clarendon, 1962.
- Maggiore, Robert. *La philosophie au jour le jour*. Paris: Flammarion, 1994.
- Mu Soeng. *The Diamond Sutra: transforming the way we perceive the world*. Boston: Wisdom, 2000.
- Rescher, Nicholas. *Paradoxes: Their Roots, Range, and Resolution*.
- Suzuki, D.T. *The Zen Doctrine of No-Mind*. Ed. Christmas Humphreys. York Beach, ME.: Weiser, 1972.
- Trungpa, Chögyam. *Illusion's Game*. Shambhala: Boston, 1994.

## Main References

### **Works by Avi Sion**

Avi Sion is the author of several works on logic and philosophy, including: *Future Logic* (1990); *Judaic Logic* (1995); *Buddhist Illogic* (2002); *Phenomenology* (2003); *The Logic of Causation* (1999, 2003, 2010); *Volition and Allied Causal Concepts* (2004); *Ruminations* (2005); *Meditations* (2006); *Logical and Spiritual Reflections* (2008-9); *A Fortiori Logic* (2013).

These works can be freely studied at:  
[www.TheLogician.net](http://www.TheLogician.net) and [www.AviSionBooks.com](http://www.AviSionBooks.com).

*Paradoxes and their Resolutions* can be read online at:  
[www.avisionparadoxes.bravesites.com/](http://www.avisionparadoxes.bravesites.com/)

A complete list of Avi Sion's websites can be viewed at:  
[www.avisionlinks.bravesites.com](http://www.avisionlinks.bravesites.com).

Avi Sion can be reached for questions or comments at:  
[avi-sion@thelogician.net](mailto:avi-sion@thelogician.net)

