Chapter 5

Paradoxes of Definability, Russell's Paradox, the Liar

I. Paradoxes of Definability

Following Richard, write in alphabetical order all permutations of pairs of letters of the alphabet ('aa', 'ab', ... 'az', 'ba', 'bb', ...), followed by all permutations of triples of letters of the alphabet taken in alphabetical order, and so on for quadruples, quintuples and so on. Cross out all permutations of letters that do not define real numbers. Let E be the resulting enumeration of permutations that define real numbers, and arrange these reals in the order in which they are defined. (For any number that is defined by more than one permutation, we choose the permutation that is first in alphabetical order.[1]) Now, consider the following collection of letters:

> Let p be the digit in the nth decimal place of the nth number defined by a member of E. Now form a number having 0 for its integral part and, in its nth decimal place, p+1 if p is not 8 or 9, and 1 otherwise.

Call the number so defined N. Then N cannot be in the enumeration of reals defined by members of E. If it were the nth number, then the digit in its nth place would be the same as the digit in the nth decimal place of the nth number, which is not the case. Yet N is defined by a permutation of letters: so it is in the enumeration of reals defined by members of E. So we have a contradiction.[2]

Next, König's paradox: consider those real numbers that can be defined by finitely many words of English. These reals form a denumerable set. Now consider the non-denumerably many reals that cannot be so defined. If these reals can be well-ordered, then there is a least member. But now, as we have just demonstrated, this 'undefinable' real *can* be defined in finitely many words. So we have a contradiction. König concluded that the reals cannot be well-ordered -- but this conclusion is unacceptable in the light of Zermelo's well-ordering theorem.[3] So the contradiction remains, and we are confronted with a paradox.

Following Russell, observe that only a finite number of English expressions can be formed from a given finite number of syllables. So there must be positive integers that can be defined only by expressions with at least thirty three syllables, and among these integers there is a least. Now consider this expression of English: "the least positive integer which is not definable by an English expression with fewer than thirty three syllables". This expression has thirty two syllables, so the number not definable in fewer than thirty three syllables is definable in thirty two syllables – and we have a contradiction. This is Berry's paradox.[4]

These three paradoxes – Richard's, König's, Berry's – are the so-called *paradoxes of definability*. This is the traditional label -- but the label may be misleading in two ways. First, the appearance of a modal element in the term 'definable' is deceptive: for example, we might just as well have presented Berry's paradox in terms of the phrase 'the least positive integer which is not defined by an expression of English containing fewer than thirty four syllables'. Second, the paradoxes do not turn on any technical sense of definition. When we generate these paradoxes, we count as a definition any phrase that denotes or refers to a number; so, for example, the phrase 'the number of moons of Mars' will count as a definition of the number 2. The paradoxes turn on the semantic relation that holds between a referring expression and its referent, whether the relation is expressed by 'defines' or 'denotes' or 'refers to'. The paradoxes would be better called paradoxes of reference, or paradoxes of denotation.

We now apply the singularity account to the paradoxes of definability.

I.2 Richard's paradox

Here is a discourse associated with the Richard paradox, presented in four segments.

(I) Scene-setting

> Obtain an enumeration E of all English phrases, arranged alphabetically and according to length. Some of these phrases will denote real numbers, and these numbers can be arranged in the order of their denoting phrases. Let the Richard phrase be the following phrase in E: "the real which has 0 for its integral part, and, in its nth decimal place, p+1 if p is not 8 or 9, and 1 otherwise, where p is the digit in the nth decimal place of the nth number denoted by a phrase in E."

(II) Uncovering pathology

> Suppose the Richard phrase denotes a real, say q. Then q is the real which has 0 for its integral part, and, in its nth decimal place, p+1 if p is not 8 or 9, and 1 otherwise, where p is the digit in the nth decimal place of the nth number denoted by a phrase in E. Now suppose that q is the kth number denoted by a phrase in E. Then the digit d in the kth place of q is distinct from the digit in the kth place of the kth number – that is, d is distinct from itself. We are landed in a contradiction. We infer:
>
> The Richard phrase is pathological, and does not denote a real.

(III) Stock-taking

> Since the Richard phrase is pathological, it is a member of E that does not denote a real number. Once it is recognized as pathological, we are left with the reals that are denoted by the other phrases in E. And so there is a number which is the real which has 0 for its integral part, and, in its nth decimal place, p+1 if p is not 8 or 9, and 1 otherwise, where p is the digit in the nth decimal place of the nth number denoted by a phrase in E.

(IV) Re-evaluation

> But here is the Richard phrase again. And so we conclude:

(Σ) The Richard phrase does denote a real.

This discourse is an instance of *Repetition*. At the second stage, a token of the Richard phrase is produced, and we reason to the conclusion that the Richard phrase is pathological and fails to denote.   At the third stage, we take stock of this pathology, and produce another token of the Richard phrase.   At the fourth stage, we conclude that the Richard phrase successfully denotes.

The earliest discussions of Richard's paradox show some sensitivity to the Repetition reasoning.   In response to his own paradox, Richard offers reasons why we should treat the Richard phrase as pathological.[5]   So Richard uncovers pathology and eliminates the Richard phrase from the enumeration of phrases that denote real numbers – we have to "cross it out".[6] Peano recognized the difficulty that this generates. If we cross out the Richard phrase, we are left with an enumeration of unproblematic phrases that denote real numbers – and given this enumeration, we can use the Richard phrase to denote a number.   Peano writes:

> "If the [Richard] phrase that defines N does not express a number, as was demonstrated above, then, when I calculate N, I pass by this phrase, which does not define a number, and the definition of N acquires a meaning. That is to say, if N does not exist, then it exists."[7]

Peano could be clearer here (for example, how could the Richard phrase "not express a number" and at the same time define N?) – but it seems clear that Peano is sensitive to the way in which the failure of the Richard phrase to denote a number provides the very grounds for its subsequent success.[8]

Although there is an obvious similarity between this instance of Repetition and the Repetition reasoning about C, P and L, there is the following difference: the Richard phrase is initially introduced as a phrase-type, not a token. We start out with the assumption that we can enumerate the expressions of English. In making this assumption we must have in mind expression-*types*. Recall Richard's recipe: we write in alphabetical order all permutations of pairs of letters of the alphabet ('aa', 'ab', ... 'az', 'ba', 'bb', ...), followed by all permutations of triples of letters of the alphabet taken in alphabetical order, and so on for quadruples, quintuples and so on. Phrases obtained in this way -- constructed via letter-types -- are individuated only via their types. So at the outset of our reasoning, we start with an enumeration of the phrase-types of English – and among these phrase-types is the Richard phrase.

But then in the second segment, the Richard phrase is *used* – a token of the Richard phrase appears in segment (II). In reasoning to the pathology of the Richard phrase, there is no avoiding its *use*. The Richard phrase is embedded in some particular context of use. To anticipate the contextual analysis: when we conclude that the Richard phrase is pathological, it is not the Richard phrase qua phrase-type that we evaluate as pathological, but rather the phrase-type *in a given context*. According to the contextual analysis, 'denotes' is a context-sensitive term. So the Richard phrase-*type* contains a context-sensitive term, and consequently can no more be evaluated independently of a context than the phrase-types 'the next person in line' and 'my favorite number'.[9] According to the singularity account, we need assume very little about the context of use of the Richard phrase in segment II: all that matters is that the context is unreflective with respect to the phrase. We can abstract away from who says it or writes it, and where and when. And I'll assume that we can talk about the denotations in this context of the phrase-types in E, since all tokens of any of these phrase-types have the same denotation, whatever the context.

Let R be the token of the Richard phrase that appears in the second sentence of segment (II). Let c_R be R's context of utterance, where the key feature of the context c_R is that it is unreflective with respect to R. In the familiar way, we will represent the occurrence of 'denotes' in R, and any coextensive occurrence, by 'denotes$_{cR}$'.   So the members of R's determination set – the members of E – are evaluated by the c_R-schema. R itself is a member of E, so R is evaluated by the c_R-schema.   So the primary representation of R is <type(R),c_R,c_R>.   R's place in the sequence E will be determined by its type, but unlike the other members of E, R is a token expression, and so it will be represented not simply by its type, but rather by its primary representation.

At the outset of segment (II), we assume that the Richard phrase denotes a real number q, and go on to identify q via the token R. So the reasoning proceeds with this token of the Richard phrase, and the following instance of the c_R-schema:

> R denotes$_{cR}$ q if and only if q is the real which has 0 for its integral part, and, in its nth decimal place, p+1 if p is not 8 or 9, and 1 otherwise, where p is the digit in the nth decimal place of the nth number denoted$_{cR}$ by a phrase in E.

We go on to reach a contradiction, since R is a diagonal definition, and a member of E. The subconclusion reached at the end of segment (II) is captured by:

> R is pathological, and does not denote$_{cR}$ a real.

The Richard phrase as it used in segment (II) – namely, R -- does not have denotation_{cR} conditions, and so it is a singularity of 'denotes_{cR}'.

Now consider the third segment. Here, we take stock of the deliverances of the c_R-schema, and as a result, we produce a repetition of R, call it R*. The contextual analysis runs as follows:

> Since the Richard phrase is pathological, it is a member of E that does not denote a real number. Once it is recognized as pathological, we are left with the reals that are denoted_{cR} by the other phrases in E. And so there is a number which is the real which has 0 for its integral part, and, in its nth decimal place, p+1 if p is not 8 or 9, and 1 otherwise, where p is the digit in the nth decimal place of the nth number denoted_{cR} by a phrase in E.

Compare the repetition R* with the repetition C*. You produce C* and identify a denotation for it when you take stock of the deliverances of the c_C-schema: C fails to denote_{cC}, and A and B denote_{cC} π and 6 respectively. In the Richard case, we identify a denotation for R*, given that R fails to denote_{cR}, and the expression-types in E denote_{cR} real numbers.

At the fourth stage, we explicitly evaluate R*. Let c_Σ be the context in which the evaluation Σ is produced, so that '$\text{denotes}_{c\Sigma}$' represents the occurrence of 'denotes' in Σ. Then R* is evaluated by the c_Σ-schema, a schema that is reflective with respect to R. At the third stage we found that R* denotes a real number, and now we accommodate that finding by employing a schema that takes into account R's pathology. That is, the c_Σ-schema is an r_R-schema.

A different version of the Richard paradox provides an instance of *Rehabilitation*. Keep (I) and (II) as before, and now continue as follows:

(III′) <u>Stock-taking</u>

> Since the Richard phrase is a pathological denoting phrase, it is not among the phrases of English that denote real numbers. Once the Richard phrase is recognized as pathological, we are left with the reals that are denoted by the other phrases in E. Take these real numbers and reconsider the words that make up the Richard phrase: these words constitute a diagonal definition of a new real number.

(IV′) <u>Re-evaluation</u>

> So,
>
> (2′) The Richard phrase does denote a real.

In this discourse, the token R that appears at stage II is re-evaluated in the light of its pathology (that is, its failure to denote$_{cR}$). (Here, Peano's words seem particularly apt: "If the [Richard] phrase that defines N does not express a number, as was demonstrated above, then, when I calculate N, I pass by this phrase, which does not define a number, and the definition of N acquires a meaning"). Taking into account its pathology, we re-examine R and find that it does denote -- once it is evaluated by an r_R-schema.

We can give more formal expression to these ideas via the primary trees for R and R*. R's primary tree is given by:

<type(R),c_R,c_R>

/ | … \ …

phrase-type phrase-type … <type(R),c_R,c_R> …

/ | … \

phrase-type phrase-type ... <type(R),c_R,c_R> …

/ | … \

phrase-type phrase-type ... .

.

.

Since R’s primary representation repeats on the infinite branch, the tree shows that R cannot be evaluated by the c_R-schema: R is a singularity of 'denotes$_{cR}$'. The tree indicates how we can reflectively determine a denotation for R: eliminate R’s primary representation at the second tier (“cross it out”, to use Richard’s words), and determine a denotation for R in terms of the other members of R’s determination set.

The primary representation of the repetition R* is <type(R),c_R,r_R>.[10] R*'s determination set is the same as R's, comprising the members of E. R's pathology is indicated in the primary tree for R*:

<type(R),c_R,r_R>

/ | … \ …

phrase-type phrase-type … <type(R),c_R,c_R> …

/ | … \

phrase-type phrase-type ... <type(R),c_R,c_R> …

/ | … .

phrase-type phrase-type ... .

.

This primary tree displays a procedure for determining a denotation for R*. We need first to determine the denotation$_{cR}$ of the members of R*'s determination set. The repetition of R's primary representation on the infinite branch indicates that R has no denotation$_{cR}$ - and so we need only determine the denotation$_{cR}$ of the other members of E. As we would expect, the denotation$_{rR}$ of R* is the same as the reflectively determined value of R.

Once we've concluded that R* denotes$_{c\Sigma}$ a real, we can form a new enumeration $E^{/}$ of the phrase-types that denote reals, together with R* (R*'s position in the enumeration will be determined by its type). Both E and $E^{/}$ can be described as an enumeration of phrase-types that denote reals, together with the Richard phrase – but E and $E^{/}$ are different enumerations because they contain different tokens of the Richard phrase-type. We could now continue on from the conclusion Σ of stage IV of *Repetition* as follows (for the intuitive reasoning, ignore the subscripts):

> Given that R* $\text{denotes}_{c\Sigma}$ a real, there is a distinct real which has 0 for its integral part, and, in its nth decimal place, p+1 if p is not 8 or 9, and 1 otherwise, where p is the digit in the nth decimal place of the number $\text{denoted}_{c\Sigma}$ by the nth phrase in $E^{/}$.

The Richardian phrase-token here – call it R** -- denotes (that is, $\text{denotes}_{c\Sigma}$) a new real number. And we could go on to obtain $E^{//}$ by add R** to $E^{/}$, and produce a new Richardian phrase-token R*** that makes reference to $E^{//}$, and denotes ($\text{denotes}_{c\Sigma}$) a new real number. And so on. In this way, we obtain a denumerable sequence of Richardian phrase-tokens, where each denotes ($\text{denotes}_{c\Sigma}$) a distinct real.[11]

This kind of iteration is benign. It can seem otherwise if we fail to separate out distinct Richardian phrases, or fail to discern changes in the determination sets. R*, R**, … are all different tokens of the same type, each with its own determination set and its own singularities. Here's a continuation of *Repetition* that slurs over these differences:

(V) Iteration

> We have just concluded that the Richard phrase denotes a real, say m. So the Richard phrase is a member of the enumeration E of expressions that denote reals. Obtain the real which has 0 for its integral part, and, in its nth decimal place, p+1 if p is not 8 or 9, and 1 otherwise, where p is the digit in the nth decimal place of the nth number denoted by a phrase in E. Here is the Richard phrase again, and given that it is a diagonal definition, it denotes a number distinct from m. So the Richard phrase both denotes m and does denote m – and we have a contradiction.

The reasoning here fails to distinguish R* and R**, and fails to distinguish their determination sets. There is no single Richardian token that denotes both m and another real; and the Richard phrase-*type*, taken in isolation from a context, fails to have a denotation, since it contains a context-sensitive term. So the contradiction associated with *Iteration* is illusory.

I.3 Berry and König

A version of the Berry discourse exhibits *Repetition*, as detailed below. Notice that, even though the Berry phrase is first presented as a phrase-type, the Berry phrase, like the Richard phrase, is *used* in the course of uncovering pathology. For economy, the discourse is presented with contextual subscripts in place - for the intuitive reasoning, ignore the subscripts. The first token B of the Berry phrase is produced at the second stage, where we uncover pathology. Let c_B be the context associated with B, so that the occurrence of 'denotes' in B (and any co-extensive occurrence) is represented by 'denotes$_{cB}$'. As with the Richard case, we need assume little about the context c_B – its key feature is that it is unreflective with respect to B, and we can abstract away speaker, time and place. A repetition B* of B occurs at the third stage. Let $c_\mathcal{E}$ be the context of the evaluation $\mathcal{E}$ of B* at the fourth stage, so that 'denotes$_{c\mathcal{E}}$' represents the occurrence of 'denotes' in $\mathcal{E}$. So B* is evaluated by the $c_\mathcal{E}$-schema, a schema reflective with respect to B – that is, an r_B-schema.

(I) Scene-setting

> Consider the English expression "the least positive integer which is not denoted by an English expression with fewer than thirty two syllables". Call this the Berry phrase.

(II) Uncovering pathology

> Suppose the Berry phrase denotes$_{cB}$ a number, say k. But the Berry phrase has fewer than thirty two syllables. So k is the least positive integer which is not denoted$_{cB}$ by an English expression with fewer than thirty two syllables - and yet k is denoted$_{cB}$ by an English expression with fewer than thirty two syllables. This

is a contradiction. We conclude that the Berry phrase is pathological, and does not denote$_{cB}$ an integer.

(III) Stock-taking

Since the Berry phrase is pathological and fails to denote$_{cB}$, we have to "cross it out". And that will leave us with the finitely many positive integers that are denoted$_{cB}$ by unproblematic expressions of English with fewer than thirty two syllables. Now consider all these integers, and let n be the least integer that is not among them. That is, n is the least positive integer which is not denoted$_{cB}$ by an English expression of English with fewer than thirty two syllables.

(IV) Re-evaluation

But here is the Berry phrase again. And so:

(Ɛ) The Berry phrase does denote$_{c\varepsilon}$ an integer.

The contextual analysis runs parallel to the Richard case. The evaluation of B by the c_B-schema leads to contradiction. B is a singularity of 'denotes$_{cB}$', as its primary tree will indicate. B* is evaluated by the reflective c_ε-schema, and this token is not pathological, as its primary tree indicates. The treatment of *rehabilitation* and *iteration* also run parallel to the Richard case.[12]

For a version of the König discourse that exhibits *repetition*, the Berry phrase is replaced by “the least ordinal not denoted by an English phrase”. The treatment of *rehabilitation* and *iteration* for the König case runs parallel to that of the Richard and the Berry.

II. Russell’s Paradox

It seems that some sets, like the set of abstract objects or the set of infinite sets, are members of themselves. Other sets, like the set of U.S. Senators or the set of bridges over the Thames, are not. What about the set of exactly those sets that are not members of themselves?

In the familiar way, we reach a contradiction if we suppose that it is a member of itself, and if we suppose that it isn't.

There is a natural assumption underlying Russell's paradox: the assumption that every predicate has an extension. Call this assumption Naive Comprehension, symbolized by

$$\exists y \forall x (x \in y \leftrightarrow \varphi x),$$

where φ stands for any predicate, and the variables range over extensions of predicates. Naïve Comprehension guarantees an extension for, say, the predicate 'bridge over the Thames'. But certain substitutions for φ lead to trouble. In particular, put the predicate '$x \notin x$' for φ, and we are landed in Russell's paradox.

It is easy to slide, as we have just done, between two distinct versions of Russell's paradox, one in terms of sets and one in terms of the extensions of predicates. In my view, sets and extensions provide very different settings for Russell's paradox. To put the difference in a rough and preliminary way, we can say that sets are 'mathematical' objects, formed by assembling or combining individuals, while extensions are 'logical' objects, arising out of predication. It is far from clear that a resolution of the paradox in one setting carries over to the other.

II.1. Sets and Classes

In Zermelo-Fraenkel set theory (ZF), Naive Comprehension is abandoned in favor of the axiom schema of *Separation*:

$$\forall z \exists y \forall x (x \in y \leftrightarrow x \in z \& \varphi x)$$

where the variables range over sets. Given a predicate φ, we are no longer guaranteed a set of elements that are φ; rather, given a set z, there is a subset y of those members of z that are φ. Zermelo placed a fundamental restriction on *Separation*: the predicate must be *definite*. Zermelo himself was only vaguely characterized by Zermelo, but was subsequently sharpened by Skolem: take the language of set theory, precisely formulated in the usual recursive way – then a definite condition is simply one expressed by a 1-place predicate of the language.[13] Clearly, versions of the definability paradoxes now do not arise, since semantic predicates such as 'defines' or 'denotes' are no part of the language of set theory. And Russell's paradox is avoided because although the predicate '$x \notin x$' is definite, there is no set of *all* the non-self-membered sets (though we can always separate off from a given set the subset of its members that are non-self-members).

So according to Zermelo, we must abandon a predicative conception of set that allows arbitrary set formation: "it no longer seems admissible today to assign to an arbitrary logically definable notion a set, or class, as its extension".[14] And in his 1908 paper, Zermelo assumed that there is no alternative conception of set that is as simple as the predicative conception *and* free of contradiction. We have to approach things differently:

> "Under these circumstances there is at this point nothing left for us to do but to proceed in the opposite direction and, starting from set theory as it is historically given, to seek out the principles required for establishing the foundations of this mathematical discipline. In solving the problem we must, on the one hand, restrict these principles sufficiently to exclude all contradictions and, on the other, take them sufficiently wide to retain all that is valuable in this theory".[15]

Zermelo adopts a pragmatic stance: the task is to find axioms that avoid the paradoxes and yield "the entire theory created by Cantor and Dedekind".[16] Zermelo's axiomatization was not driven by any particular conception of set.

Yet such a conception is available, as Zermelo and others subsequently saw.[17] It is the *combinatorial* or *iterative* conception.[18] Think of a set as formed this way: we start with some individuals, and collect them together to form a set. Suppose we start with individuals at the lowest level. At the next level, we form sets of all possible combinations of these individuals. And then we *iterate* this procedure: at the next level, we form all possible sets of sets and individuals from the first two levels. And so on.

In pure set theory we start with no individuals, just the empty set, the existence of which is given by the Empty Set Axiom. Other set existence axioms assure us that we can build new sets out of old ones. For example, the Pairing Axiom tells us that, given sets A and B, there is a set having as its members just A and B; the Union Axiom tells us that for any set A, there exists a set whose members are exactly the members of the members of A; and the Power Set Axiom tells us that, given any set A, there exists a set whose members are exactly the subsets of A. Given the set of all sets at a particular level, the next level will contain the members of its power set - we can think of the Power Set Axiom as a driving force behind this endless cumulative hierarchy. Every set appears somewhere in this hierarchy.[19] On the combinatorial/iterative conception Russell's paradox is avoided because there is no set of exactly the non-self-membered sets. According to the combinatorial/iterative conception, no set is a member of itself. (No set can collect or 'lasso' itself.)[20] So the Russell set, if it existed, would be the universal set. But there is no universal set in the iterative hierarchy.

Now arguably ZF provides a suitable foundational theory for mathematics.[21] But we may wonder if we have a satisfactory resolution of Russell's paradox. For one thing, we expect a well-defined predicate to have an extension. In particular, we expect the self-identity predicate to have an extension - but none of the sets in the ZF hierarchy is the extension of 'x=x'. Similarly with the predicate '$x \notin x$'. Or again, since ZF provides a clearcut concept of *set*, we expect the predicate 'set' to have an extension - and in ZF it doesn't. And this may strike us as *ad hoc*; with Frege, we can ask "… how do we recognize the exceptional cases? "[22]

Further, in ZF we quantify over sets, and so we need a domain of quantification – but no set in the hierarchy can serve as this domain. In his 1908 axiomatization of set theory, Zermelo introduces a domain B of individuals, "which we shall call simply *objects* and among which are the *sets*."[23] He goes on to prove that the domain B "is not itself a set", and goes on to say that "this disposes of the Russell antinomy so far as we are concerned."[24] So the set/non-set distinction was prominent from the very beginnings of axiomatic set theory. How is this domain given to us? Clearly not by the ZF axioms, or via the iterative conception. It's given to us predicatively, as the collection of those things that are either urelements or sets – as the extension of the predicate 'is an urelement or a set'. Zermelo simply assumes that the domain exists, and implicitly assumes that it can be specified predicatively. But now we have a new kind of collection in the picture. As Skolem puts it: "If we adopt Zermelo's axiomatization, we must, strictly speaking, have a general notion of domains in order to provide a foundation for set theory".[25] It seem that s*ets* are not enough – *extensions* are needed too. We have two distinct conceptions in the picture: the iterative conception of ZF sets, and the predicative conception of extensions. And if extensions are different in kind from sets, we need a separate treatment of Russell's paradox for extensions.

In later work, Zermelo offered a way of resisting this line of thought. He argued that the set/non-set distinction is a *relative* one. In a 1930 paper, in contrast to his earlier pragmatic line, Zermele argues that the ZF axiomatization provides a compelling explanation of the paradoxes. He shows that the 'general' axioms of set theory[26] are not categorical – they have non-isomorphic models. Indeed, they are satisfied by an unending sequence of models, each one associated with a distinct ordinal number – what Zermelo calls the 'boundary number' -- which indexes the level of the model in a cumulative hierarchy. The universe or domain of any one of these models will not be a set of the model, on pain of paradox. But the non-sets of one model will be sets of the models further along in the sequence. Consider the domain of a given model. This is paradox-producing only if we mistakenly regard it as a set of the model, when in fact it is a non-set of the model. And it is easy to make this mistake if we think of the model as absolute, as containing all sets, as comprising "set theory itself".[27] But the model is just one of an unending series, a series that "reaches no true completion in its unrestricted advance, but possesses only relative stopping points".[28] The domain of a given model is both a non-set and a set, but there is no contradiction – it's a non-set of the model for which it serves as domain, but a set of all higher models. And similarly for the paradoxical 'sets', like Russell's: they are non-sets in a given model, but sets in subsequent higher models. The set/domain distinction is relative.[29]

How successful is Zermelo's resolution? One problem is this: while the axiom of Separation allows the predicative formation of sets in a restricted way (you must start with a set), the domains are formed predicatively in a way that goes beyond Separation. Domains are not subsets of a given set, but are formed directly from the predicate 'is an urelement or a set'. As Michael Hallett puts it, "what we have is a return to something like a principle of arbitrary set formation, a principle which Zermelo repudiates in his axiomatization of 1908…".[30]

Further, consider the predicate 'is a boundary number'. This expresses a notion in terms of which the series of ZF models is described. We've just noted that Zermelo permits the formation of collections from predicates unconstrained by Separation. So we have as yet no reason to deny that the predicate 'is a boundary number' determines a collection. And a relativized treatment of the extension of this predicate is not possible, on pain of failing to describe the intended series of models. So here we have a non-set that cannot also be a set, even in a relativized way.[31] And so we can now ask whether Russell's paradox arises for this kind of collection. Until we have an answer to this question, we have not yet disposed of Russell's paradox.

We should also put pressure on Zermelo's rejection of an *absolute* universe of sets. Suppose I am asked about the range of my set-theoretic quantifier – then I might well reply that it *cannot* be a set, and that there is no more inclusive universe in which it is a set. If I study more set theory, and learn about, say, Zermelo's boundary numbers and inaccessible cardinals, then I have come to a better understanding of the universe of sets – but why suppose that my universe has expanded? D.A. Martin writes: 'Though in some sense my coming to hold that inaccessibles exist has perhaps changed my concept of set, the notion of all-inclusiveness, of absolute infinity, was already a central part of my concept".[32]

So if we are unpersuaded by Zermelo's relativistic approach, the question remains: Can a set theory by itself accommodate the intuitions that drive us towards non-sets, in particular extensions. At this point, it is natural to consider set theories with a universal set. Such a set theory will provide a *set* to serve as the extension of 'x=x'. An example is Quine's NF, in which we can prove the existence of a universal set V.[33] However, Russell's paradox is avoided because there is no set of all the non-self-membered sets – there is no set serving as the extension of the predicate '$x \notin x$'.[34] The same is true of the predicate 'x is well-founded', on pain of Mirimanoff's paradox.[35] The systems of Church and Mitchell - two other set theories with a universal set - suffer similar limitations.[36] In response to these limitations, we might admit subcollections of V (for example, the collection of non-self-membered sets, or the collection of well-founded sets) that are not sets. Here one might draw on Vopenka and Hayek's notion of a semiset.[37] A semiset is a subclass of a set, and a proper semiset is a semiset that is not itself a set.[38] So, in the context of a set theory with a universal set, we might admit as semisets the collections given by the predicates '$x \notin x$' and 'x is well-founded', where these semisets are subclasses of V that are not themselves sets.[39] But semisets are given via predication.[40] So we have not escaped non-sets that are predicatively given.

A more familiar way to distinguish sets and non-sets is by the introduction of proper classes. There is a class of all sets, a class of all non-self-membered sets, a class of all ordinals, but no such sets. These *proper classes*, are, in von Neumann's phrase, "too big" to be sets.[41] For von Neumann, all sets are classes, but not all classes are sets.[42] Further, these proper classes cannot themselves be members.[43] And so a proper class of all sets, or of all non-self-membered sets, do not generate paradox. A further assumption is needed, that these proper classes can themselves be members (and in particular members of themselves), and this assumption is false.

We can formulate an Axiom of Comprehension for classes as follows:

$$\exists A \forall x (x \in A \leftrightarrow \varphi x),$$

where A is a variable ranging over classes, x is a variable ranging over sets, and there is the following restriction on : φ does not contain quantifiers over classes. φ can contain class variables (or parameters), but these must not be bound.[44] Comprehension for classes guarantees the existence of a class of all sets (put 'x=x' for φ) and the existence of a class of exactly the non-self-membered sets (put 'x∉x' for φ). And similarly there is a class of all ordinals and a class of the well-founded sets. And in von Neumann's system, the arguments that led to paradox instead establish that these classes are proper classes. And we cannot generate a new family of paradoxes for classes. No proper class can be a member, and so no class can have a proper class as a member. In particular, there can be no class of all classes, and no class of all the non-self-membered classes.

In von Neumann's system, then, there are extensions for those definite predicates - like 'x=x', 'x∉x', 'well-founded set', 'ordinal' - which cannot be assigned a set as their extension; there is a well-determined collection of all the ZF sets; and there is a domain for quantification over sets. But the intuitive costs are very high.[45] First, it seems *ad hoc* and counterintutive to say that a proper class cannot be a member. Why, for example, can't we form its unit class? More generally, the ban on finite classes of proper classes is not well-motivated - finite classes do not generate paradox.[46] Second, we expect the predicate 'class' to have an extension, just as we expect the predicate 'set' to have an extension. Moreover, von Neumann quantifies over proper classes. In von Neumann's terminology the proper classes correspond to the II-objects that are not I-objects, and throughout his 1925 paper, von Neumann quantifies over II-objects. His arithmetic construction axioms and logical construction axioms are ways of producing II-objects, and they take this form: "There is a II-object such that ...".[47] We seem to need a domain for such quantification, but in von Neumann's system there is no class or extension of all classes. The problem has just been pushed back: we are left again with a clearcut concept - now, the concept of class - which has no extension.

Von Neumann's system has been liberalized in various ways. A natural move is to ease the restriction on φ in class comprehension by admitting class quantifiers, a move that has been made by Wang, Morse and Kelley.[48] In this stronger system, we can prove the existence of more sets and more classes of sets. But still a class cannot be a member - again we cannot even form the unit class of a proper class. And still there is no extension for the predicate 'class'.

One can liberalize further, and develop systems in which proper classes are members. Following Levy *et al*,[49] we can introduce the notion of a hyperclass, via the schema:

There is a hyperclass whose members are all classes (and sets) that are φ,

where φ is unrestricted.[50] Here, proper classes are members of hyperclasses (though no class or hyperclass is a member of itself). And we can go further still. Levy *et al* present a two-tier theory, with sets in the lower tier and classes in the upper tier.[51] In the upper tier, we find the class V of all sets, and its power class *P*V, and *PP*V, and so on.[52] This system can be naturally modelled in ZF#, where ZF# is the system of ZF with the additional axiom "There is at least one inaccessible cardinal". If κ is a fixed inaccessible cardinal, we may interpret our two-tier theory as follows: a 'set' is a member of rank(κ), and a "class" is a set.

And we need not stop here. We can develop a theory of classes modelled by ZF plus an axiom that asserts the existence of arbitrarily large inaccessible cardinals - a system of tiers, with the lowest occupied by sets, the next by classes, the next by superclasses, and so on. Such a theory provides a universe for category theory.[53]

Since such systems can be constructed without falling foul of the paradoxes, the ban on membership for proper classes seems not just counterintuitive, but also an unnecessarily heavy-handed response to the paradoxes. We need only a ban on *self-membership*, and more generally on *unfounded* classes - and the ban is preserved even in these liberalized systems. And with this ban we still pay the same high price: whether the system contains proper classes, hyperclasses or a series of superclasses, still there will be no collection of them all, no domain for quantification over them, and no extension for the predicate 'class', 'hyperclass' or 'superclass'.

There is a set theory in which we find self-membership and unfounded sets: Aczel's set theory, where ZF's Axiom of Regularity is replaced by an Anti-Foundation Axiom.[54] Despite this departure from ZF, Aczel's system deals with the paradoxes in just the same way as ZF: for example, there is no set of the non-self-membered sets, no set of all sets, and no set of the wellfounded sets, so Russell's paradox, Cantor's paradox and Mirimanoff's paradox do not arise. Aczel's set theory fares no better than ZF when it comes to providing extensions for the predicates 'non-self-membered set', 'set', 'well-founded set', and so on.

II.2 Extensions

We have discussed a wide variety of set theories and class theories. Setting aside their differences, there is one persistent failing that they share -- they fail to provide extensions for certain predicates. They fail to do justice to a conception of set quite different in spirit from the iterative/combinatorial conception. The alternative conception goes like this: to any predicate that denotes a well-determined condition, there corresponds the collection of those things to which the predicate applies. Corresponding to the predicates 'abstract', 'set', and 'class', for example, there are the collections of abstract things, of sets, of classes. Call this the *predicative* conception, and call the collection of things to which a given predicate applies the *extension* of the predicate. We can also put things in Fregean terms: every concept (in Frege's sense) has an extension.For Frege, the extension of a concept is a *logical* object. Frege writes: "By means of our logical faculties we lay hold upon the extension of a concept, by starting out from the concept."[55] And it is clear that Frege would be quite opposed to the iterative or combinatorial conception of set embodied in ZF. Frege writes:

> "I do, in fact, maintain that the concept is logically prior to its extension; and I regard as futile the attempt to take the extension of a concept as a class, and make it rest, not on the concept, but on single things."[56]

In a number of places, Frege argues more specifically that the combinatorial conception does not support the notion of the empty class,[57] or the notion of an infinite class.[58]

For Frege, a legitimate class theory is a theory of logical objects: a theory of extensions tied logically to concepts and predicates. It is not a theory of collections assembled arbitrarily from individuals. Russell's view of the matter is similar:

> "A class or collection may be defined in two ways that at first seem quite distinct. We may enumerate its members, as when we say, "The collection I mean is Brown, Jones, and Robinson." Or we may mention a defining property, as when we speak of "mankind" or "the inhabitants of London". The definition which enumerates is called a definition by "extension", and the one which mentions a defining property is called a definition by "intension". Of these two kinds of definition, the one by intension is logically more fundamental."[59]

Or as Forster puts it:

> "... set membership can be seen as an allegory for predication, so that sets are deemed to arise as extensions of predicates."[60]

According to this idea, Forster says, "[s]et theory is a branch of *logic*, the logic of *predication*."[61]

So we now have in play three types of collections: sets, classes, and extensions. Is any one of these primary? Is there one to which the others can be reduced? Are there distinct intuitions underlying each type? There are clearly different conceptions underlying ZF sets and extensions: we can take ZF sets as collections conceived of iteratively or combinatorially, and extensions as collections conceived of predicatively. Does one reduce to the other? [62]

We cannot reduce extensions to ZF sets: as we have seen, there are predicates of ZF (like '$x=x$' and '$x\notin x$') whose extensions cannot be sets. And even if, for the moment, we think of proper classes as glorified sets, still extensions won't be reducible to sets - no system incorporating proper classes provides an extension for 'class' or 'non-self-membered class'. Quine's NF does provide a *set* as the extension of '$x=x$', but we have seen that there remain other predicates (like '$x\notin x$' or 'x is well-founded') whose extensions cannot be sets of NF.

Moreover, it is natural to suppose that some extensions are members of themselves - for example, the extension of the predicate 'abstract' is itself abstract, and the extension of 'infinite extension' is itself infinite. But, as we saw, no set of ZF is a self-member - Foundation is a part of standard set theory.[63] And, as far as I am aware, there is no system with classes as well as sets that allows self-membered proper classes.

On the other hand, we cannot reduce sets to extensions. We saw that the iterative hierarchy is generated by the power set operation. And the assumption that, given a set, there exists a set of all its subsets, is not at all underwritten by the notion of predication: we do not require a predicate or a rule for determining each subset. Indeed, on certain natural assumptions, we have to give up the idea that every set is determined by a predicate. For if we assume that in a given language there are denumerably many predicates, then the members of a nondenumerable set will outrun the predicates. More generally, if we suppose that the predicates of a language always form a set of some cardinality, then, since there will always exist sets of greater cardinality, sets will always outrun predicates.

In this regard, consider also the Axiom of Choice. When in 1904 Zermelo proved the Well-Ordering Theorem (that every set can be well-ordered), he made fully explicit his reliance on the Axiom of Choice, "the principle that even for an infinite totality of [non-empty] sets there always exist mappings by which each set corresponds to one of its elements".[64] We can think of these mappings as functions that 'choose' an element from each set in the totality, yielding a 'choice set' as output. Zermelo's proof met with immediate opposition because it assumed the existence of mappings and choice sets without defining them.[65]

However, if we adopt the iterative conception such worries about the Axiom of Choice seem quite misplaced. Consider a totality of non-empty sets. There will be some level of the iterative hierarchy at which all these non-empty sets appear, and at this same level all the associated choice sets will also appear, since their members all appear at lower levels.[66] Again sets outrun the expressive capacity of our language.[67]

So I think we should regard extensions and sets as independent notions. They embody quite different ways of thinking about collections. We might regard them as primitive notions, or perhaps as alternative, mutually irreducible conceptions of the more general notion of *collection*.

How do things stand with classes, including proper classes? Both conceptions, the predicative and the iterative, provide motivation for classes. But there seems to be no distinct conception peculiar to classes. On the predicative side, we may introduce classes to serve as the extensions of Zermelo's definite predicates, with proper classes serving as the extensions of the paradox-producing definite predicates.[68] But as we have seen, not even proper classes can serve as the extension of certain predicates (e.g. 'class', or in extended systems, 'hyperclass' or 'superclass').

On the iterative side, we find a strong pull towards treating proper classes as we treat sets. Von Neumann writes:

> “... If we make the sets that are "too big" and incapable of being arguments capable of being arguments in a new system P, we can still circumvent the antinomies if in turn we admit the sets that are formed from all of these and are "still bigger" (that is too big on P) but declare them incapable of being arguments.”[69]

For von Neumann, proper classes are much like sets, except that they cannot be members. So it is natural to iterate, and expand the system containing proper classes to a system P in which they are members of still bigger classes which, in the system P, cannot themselves be members. And we can keep going in this way. This should remind us of Levy's discussion of hyperclasses, the two-tier theory (where the upper tier contains V, *P*V, *PP*V, ...), and the system of iterated superclasses. These systems treat proper classes in the same iterative way that ZF treats sets. Proper classes now look a lot like sets, occupying sufficiently high levels of the iterative hierarchy. This way of looking at proper classes is all but irresistible when we recall that the two-tier theory can be modelled in $ZF^{\#}$, and that the system of superclasses can be modelled in ZF+"there are arbitrarily large inaccessible cardinals".[70] View in this way, as an extra layer or series of layers of sets, proper classes are reducible to sets.

It would seem that classes have no place to call their own; the notion of a class cannot survive as a primitive, independently motivated concept. If we wish to develop the predicative conception, we are well-advised to develop a theory of extensions directly, since classes cannot do the job. And if we are working with the iterative conception, then we do best to regard proper classes as additional sets. Either way, the notion of a proper class is squeezed out.

II.3. Extensions and Paradox

If I am right that the notions of *set* and *extension* are independent and mutually irreducible, then there are two quite different settings for Russell's paradox. In one of its settings, Russell's paradox arises for sets. And the iterative conception handles Russell's paradox along familiar lines. Since sets are generated iteratively, there is no universal set. And since no set is a self-member, there is no Russell set of non-self-members because there is no universal set.

But the parallel breaks down in the case of extensions. Even if there is no universal extension, we must admit self-membered extensions (like the extension of 'infinite extension') as well as non-self-membered extensions (like the extension of 'teaspoon'). And a paradox is generated when we ask whether the extension of 'non-self-membered extension' is self-membered or not. We need a way out of Russell's paradox for extensions. And since extensions are not reducible to sets (or proper classes), our best strategy is to turn away from set theory and develop an independent account of extensions, one that is not subject to paradox.

Here is a Repetition discourse associated with Russell's paradox, where we consider the 1-place predicates of, say, English.

(I) Scene setting

> Among the predicates of English there are those, like 'abstract' that have self-membered extensions. And there are others, like 'teaspoon', that have non-self-membered extensions. Now consider the Russell predicate 'non-self-membered extension'.

(II) Uncovering pathology

> Suppose the Russell predicate has an extension E, so that for all x, x is in E if and only if x is a non-self-membered extension. Now suppose E is a self-membered extension – then it is non-self-membered. And suppose E is a non-self-membered extension – then it is self-membered. Either way, we reach a contradiction. We infer:
>
> The Russell predicate is pathological, and does not have an extension.

(III) Taking stock

> Since the Russell predicate has no extension, it is not among the predicates of English with well-determined extensions. Once we set it aside (along with any other related pathological predicates), we will be left with just those predicates of English which have a well-determined extension. And among these predicates will be those that have a well-determined non-self-membered extension$_{cQ}$.

(IV) Re-evaluation

> But here is the Russell phrase again. We conclude:
>
> The Russell predicate does have an extension.

In parallel with the Repetition discourses associated with the definability paradoxes, the scene-setting is in terms of predicate types, such as 'abstract', 'teaspoon', and the Russell predicate type 'non-self-membered extension'. But when we uncover pathology, we *use* the Russell predicate. We find that the Russell predicate fails to have an extension only through uses of the predicate in the course of segment II. The first token of the Russell predicate type occurs when we say “for all x, x is a member of E if and only if x is a non-self-membered extension_{cQ}”. Call this token Q, and let c_Q be the context in which Q occurs. In the usual way, we’ll let ‘extension_{cQ}’ represent the occurrence of ‘extension’ in Q, and any coextensive use of ‘extension’. As in the case of the Richard paradox, we need assume very little about the context c_Q: we can abstract away from who says it or writes it, where and when, and so on. What is crucial about the context c_Q is that it is unreflective with respect to Q. In parallel with all the other cases of Repetition we have considered, the subsequent occurrences of ‘extension’ in segments II and III will share the same extension as the occurrence in Q – so they are all represented by ‘extension_{cQ}’.

At the end of segment II we uncover pathology because we evaluate Q by the unreflective c_Q-schema, that is, the schema: *x is in the* $\textit{extension}_{cQ}$ *of ‘φ’ iff x is φ*. We suppose first that E, the extension_{cQ} of Q, is a self-membered extension_{cQ} – that is, is in the extension_{cQ} of Q. If it is, then by the c_Q-schema, E is a non-self-membered extension_{cQ}. Contradiction. We suppose second that E is a non-self-membered extension_{cQ}, And now, by the c_Q-schema, it follows that E is in the extension_{cQ} of Q – but E is the extension_{cQ} of Q, so is a self-membered extension_{cQ}. Contradiction again.

The contextual analysis of *stock-taking* runs as follows:

> Since the Russell predicate has no $extension_{cQ}$, it is not among the predicates of English with well-determined $extensions_{cQ}$. Once we set it aside (along with any other related pathological predicates), we will be left with just those predicates of English which have a well-determined $extension_{cQ}$. And among these predicates will be those that have a well-determined non-self-membered $extension_{cQ}$.

We have just produced a *repetition* of Q, call it Q*. And Q* has a well-determined extension, because problematic predicates like Q have been set aside. But Q* does not have an $extension_{cQ}$, any more than Q does. But Q* does have an extension.

At the fourth stage, when we re-evaluate the Russell phrase, it is Q* that we are evaluating. Let c_E be the context of this evaluation – then '$extension_{cE}$' represents the occurrence of 'extension' in the evaluation. The evaluating schema for Q* is the c_E-schema. This is a schema reflective with respect to Q, an r_Q-schema – a schema that evaluates Q* in the light of Q's pathology.

We can capture the reasoning in terms of primary trees. The members of the determination set of Q are the predicate-types of English and Q itself. The primary representation of Q is the triple $<type(Q),c_Q,c_Q>$. The primary tree for Q has an infinite branch on which the primary representation of Q repeats. This tree indicates that Q is pathological, and a singularity of '$extension_{cQ}$'. It is easy to check that, in contrast, the primary representation of Q* does not repeat on an infinite branch of Q*'s primary tree: Q* is not pathological, and is not a singularity of '$extension_{cE}$'.

The treatments of *Rehabilitation* and *Iteration* run parallel to the cases of the definability paradoxes. Once we recognize Q as pathological, and consider only the unproblematic predicates, we can reconsider and *rehabilitate* Q – Q has an extension when evaluated by a reflective r_Q-schema. And once we've produced Q*, we can extend the Repetition discourse, as we did in the Richard case. We can produce a new Russellian token – call it Q** -- that is represented by ‘non-self-membered extension$_{cE}$’, and whose determination set comprises the unproblematic predicate-types together with Q*. And once we’ve produced Q**, we can extend the discourse still further, and produce a sequence of Russell tokens, each with its own determination set and its own singularities. But this kind of iteration is benign.

As we saw in Chapter 4, once we have singularities on board, we can state minimally restricted principles for denotation, extension and truth. In the case of extension, this suggests an axiomatic approach to extensions, in keeping with the aim of developing an extension theory in its own right, independent of set theory. The theory I have in mind is composed of an axiom and an axiom schema. The axiom is the axiom of Extensionality, according to which two extensions are identical only if they have the same members. The axiom schema is a minimally restricted version of Comprehension. We have:

<u>Axiom of Extensionality</u>

For all extensions x, y, if x=y then, for any object z, z is in x iff z is in y.

<u>Axiom Schema of Comprehension</u>

For any context α,

(i) If φ is not a singularity of 'extension$_{\alpha}$', then for all x, x is in the extension$_{\alpha}$ of φ iff x is Φ, and

(ii) if φ is a singularity of 'extension$_{\alpha}$', then φ has no extension$_{\alpha}$

where x ranges over objects, Φ is replaced by a predicate, and φ is replaced by a name of that predicate.

Of course, a proper articulation of the axiomatic theory requires a rigorous characterization of the notion of *singularity* - and that is a primary aim of the next chapter.

III. The Liar Paradox

The phenomena of *repetition*, *rehabilitation* and *iteration* have rarely, if ever, been discussed in the cases of denotation and extension. But with truth, things are different. There has been a good deal of attention paid to what is sometimes called the *strengthened liar*, versions of which are closely related to *repetition*, *rehabilitation* and *iteration*. Strengthened liar discourses have provided a major motivation for contextual accounts of truth, though objections have been raised to contextual treatments of the strengthened liar, by, for example, Gauker and Field.[71] I turn to some of this discussion now, since it may serve to clarify my contextual analysis.

III.1 The Strengthened Liar

There are different versions and different diagnoses of the strengthened liar, and not all are related to my contextual treatment of *repetition* and *rehabilitation*. We start with the Liar sentence L, written on the board in room 213:

(L) The sentence written on the board in room 213 is not true.

Consider a strengthened liar discourse as follows:

(1) L is pathological. (From the usual reasoning about L.)

So,

(2) L is not true.

Applying the truth schema to (2),

(3) (2) is true. (From (2) and the biconditional *(2) is true iff L is not true*)

Since L, like (2), says that L is not true, it follows that

(4) L is true.

From (2) and (4), we obtain

(5) L is both true and not true.

Now suppose we allow truth value gaps, and admit sentences outside both the extension and anti-extension of 'true'. Such sentences will include pathological sentences such as L. And suppose we interpret 'not true' in L as having an extension equal to the antiextension of 'true', and we say that L is neither in the extension nor antiextension of 'true'. Then, following Gauker,[72] we can easily resist the reasoning by identifying an equivocation. Consider 'not true' as it occurs in (2). Suppose that we interpret 'not true' in (2) as we have interpreted 'not true' in L. Then (2) says that L is in the antiextension of 'true', which is at odds with (1), which says that L is not in the extension or the antiextension of 'true'. So suppose instead that we interpret 'not true' in (2) as coextensive with 'outside of the extension of "true"' (that is, 'either in the antiextension of "true" or neither in the extension nor the antiextension of "true"'). Then the occurrences of 'not true' in L and (2) are differently interpreted. But then we can no longer assume that L and (2) say the same thing. So we cannot make the move from (3) to (4). Notice that on this second interpretation, we do not have a genuine case of *repetition*: (2) is not a genuine repetition of L.

As Gauker points out,[73] we can draw a lesson from this diagnosis of the strengthened liar: the 'gap' approach is inadequate, for we can go back to the liar sentence and interpret 'not true' from the outset as applying to everything outside the extension of 'true'. And now it won't help to appeal to gaps: if we assume that L is gappy, then it's not true, and that's what it says. In my terms, L is pathological because it cannot be assessed by its associated c_L-schema. (It is the case that L is neither $true_{cL}$ nor $false_{cL}$: either assumption, that L is true or that L is $false_{cL}$, leads to contradiction. But identifying L as gappy in this way does not stop us producing a genuine repetition of L.) Since L cannot be assessed by its associated c_L-schema, it is not $true_{cL}$ – and here we have a genuine repetition. As we saw in Chapter 2, L is indeed not $true_{cL}$, because if it were, it would be assessable by the $true_{cL}$-schema, and then we'd be landed in contradiction. And in a suitably reflective context we can say that L is not $true_{cL}$, since this leads to contradiction only if the c_L-schema is available, and at the *stock-taking* stage we have rejected the assumption that L can be assessed by the c_L-schema. And so the corresponding strengthened discourse will not equivocate on 'not true', and a genuine repetition of the liar sentence will be produced. The present version of the strengthened liar contains a genuine case of *repetition*, and this is the kind of discourse that requires a contextual analysis.

Gauker goes on to question this contextual analysis:

> "The claim is that in order to avoid a contradiction, we must recognize that in any sentence ascribing truth, there is an implicit reference to a context. Well, if there is no mistake in the strengthened liar reasoning when truth is understood as relative to a context, and the only mistake is our mistake in interpreting the conclusion as a contradiction when really it is not, then we should also find no mistake when we make the relativity to context explicit…"[74]

According to my account from Chapter 2, the strengthened reasoning about L looks like this, with the relativity to context made explicit:

(1) L is pathological. (Since L cannot be assessed by its associated schema, the c_L-schema.)

(L*) L is not $true_{cL}$. (From (1). L* is a genuine repetition of L.)

At this point, we have to accommodate the truth of L*, by recognizing a new standard of evaluation – the reflective c_Σ-schema. We infer:

(Σ) L* is $true_{c\Sigma}$. (From (2), and the biconditional *L* is* $true_{c\Sigma}$ *iff L is not* $true_{cL}$.)

Since L says just what L* says,

L is $true_{c\Sigma}$.

From this and L*, we obtain

L is both $\text{true}_{c\Sigma}$ and not true_{cL}. (Compare the treatment of *iteration* in Chapter 2.)

Gauker objects to the move from L* to Σ, arguing that it isn't an instance of Semantic Ascent. He writes:

> "The argument from lines [(2) to (3)] in the original strengthened liar reasoning might have been valid by the principle of Semantic Ascent, according to which any sentence *s* implies a sentence of the form *'s' is true*. However, no such inference rule recommends the reasoning from [L* to (3)] in the reconstruction. This move is certainly not an instance of Semantic Ascent."[75]

The worry seems to be that we have no reason to ascend to a claim involving a use of '$\text{true}_{c\Sigma}$' in particular, rather than some other use of 'true' tied to some other context. But once it is understood that the repetition L* is evaluated by the reflective c_Σ-schema, then the move *is* an instance of Semantic Ascent, an instance of the move from *s* to '*s' is* $\textit{true}_{c\Sigma}$. So there is no mistake in the reasoning here. Notice also that there is no equivocation over 'not true': L and L* say exactly the same thing, that L is not true_{cL}.[76]

Gauker has a further objection that runs as follows. Since L and L* do say the same thing, yet differ in semantic status, then what each says isn't sufficient on its own to establish a truth value in its context of utterance. Gauker writes:

> "But if what is said by a token does not all by itself determine whether it is true in its own context, then, a fortiori, what is said by a token does not determine whether it is true in some other context. So from the fact that two tokens say the same and the second is true in its context, we cannot draw any conclusions about which contexts the first might be true in."[77]

So from the fact that L and L* say the same, and what L* says is true in its reflective context of use, we cannot draw the conclusion that L is true in that same reflective context of use.

But I think we can draw this conclusion. According to my analysis, the sense in which a sentence is true in its context is this: it is true when evaluated by the schema fixed by its context of use. That is why, when we attend to the strengthened reasoning, we represent L and L* by their primary representations – that's how we establish their semantic status. L's status is pathological because its evaluating schema is the c_L-schema, and L cannot be assessed by that schema (as its primary tree indicates). In contrast, L* is evaluated by its evaluating schema -- the c_Σ-schema – as true$_{c\Sigma}$. The question now is: why may we conclude that since L* is evaluated as true$_{c\Sigma}$ by the c_Σ-schema, L is too? The answer is that L* is an exact repetition of L, and so they say the same thing in a very strong sense. *Any* schema will evaluate them in exactly the same way. In the more formal terms of Chapter 4, the primary representation of L* -- <type(L),c_L,c_Σ> -- is a secondary representation of L. And the respective primary and secondary trees are identical, indicating that both L and L* are evaluated as true$_{c\Sigma}$ by the c_Σ-schema.[78]

Field considers a stretch of strengthened reasoning that in essence starts out with a liar sentence

(L) L is not true,

moves to

(2) L is not true (because L is defective)

and from there to

(3) (2) is true.[79]

Field suggests that someone who knows that L is defective, "might be inclined, if he were a gap theorist" to produce (2).[80] A gap theorist who took L to be neither true nor false might well conclude that L was not true. But we don't need to be gap theorists to be inclined to produce (2). Consider a genuine instance of *repetition*, where (2) is L*, a genuine repetition of L. Here, the extension of 'not true' in L (and in L*) is everything outside of the extension of 'true', so that the appeal to gaps has no purchase. We take pathology to consist in the breakdown of the application of the true$_{cL}$-schema to L. L is not true$_{cL}$ – if it were, the c_L-schema would apply, and we'd reach a contradiction. And exactly because the c_L-schema is no longer in play, we can say that L is not true$_{cL}$ without landing in contradiction.

Field goes on to consider two possible locations for the contextual shift in the extension of 'true'. One possibility is a contextual shift in 'true' between its occurrence in L and its occurrence in (2), so that it is only appropriate to call L "not true" after a contextual shift in 'true'. As Field points out, this would leave the truth status of L unsettled, in the sense of 'true' as used in L. Field writes:

> "If one were to say that [L] isn't true in that sense either, that would undermine the rationale for saying that there has been a conceptual shift in the move from [L] to (2). Perhaps the view is that it is somehow illegitimate to ask the truth status of [L] in the sense of 'true' used in [L]? I'm not sure how this could be argued."[81]

But according to my treatment of *repetition*, this is not the place to locate a shift in the extension of 'true'. The contextual shift that occurs is not a change in the extension of 'true', but rather a shift in the background schema of assessment. When we produce the repetition L*, we have determined that the c_L-schema cannot be applied to L, and so L is indeed not true$_{cL}$. We have moved to a reflective context where the going truth-schema is reflective with respect to L. So it's not illegitimate to ask about the truth$_{cL}$ status of L – and the answer is that L cannot be evaluated by the c_L-schema, on pain of contradiction.

The second possible location Field considers for the contextual shift in 'true' is in the step from (2) to (3). Field writes that on this view, the use of 'true' in (2) "accords with a gap theory", while the use of 'true' is (3) is a broader use of 'true'.[82] As an account of my analysis of *repetition*, where (2) is the genuine repetition L*, what Field says here is partly right and partly wrong. Partly right because we shift from ‘true$_{cL}$’ in L* to ‘true$_{c\Sigma}$’ in (3) – when we produce (3), we assess L* by the reflective cΣ-schema; partly wrong, because on the singularity account, ‘truecΣ’ is neither broader nor narrower that ‘truecL’ – each has singularities that the other doesn’t. Also partly wrong, because according to my analysis of *repetition*, the sense of ‘not true’ used in L leaves no room to escape the paradox by declaring L to be gappy, because being gappy is one way of being not true in that sense. So the use of ‘true’ in L* does not accord with a gap theory. The discourse that motivates my contextual analysis involves genuine repetition: ‘not true’ in both L and L* applies to the sentences outside the extension of ‘true’ – and that includes the gappy sentences.

Field goes on to say that the contextual view under discussion

> “would have to agree that [L*] isn’t true in the sense of ‘true’ used in [L*]; but perhaps the contextual pressures on our ordinary use of ‘true’ makes it difficult to say so.”[83]

According to my analysis of *repetition*, it is indeed right to say that L* isn’t true in the sense of ‘true’ used in L*: L* isn’t true$_{cL}$, any more than L is. And it is right to say that in the course of the repetition discourse, L* is not evaluated this way. When L* is evaluated, the c_L-schema is no longer in play. Rather, the schema fixed by L*’s context is the reflective c_Σ-schema, and this is the schema by which L* is evaluated – as true$_{c\Sigma}$. So there’s a sense in which the “contextual pressures” at work in the repetition discourse lead to an evaluation of L* by the c_Σ-schema, not the c_L-schema. However, from the point of view of the singularity theory, the evaluation of L* as not true$_{cL}$ is readily accommodated. The triple <type(L),c_L,c_L> is a secondary representation of L*, and the corresponding secondary tree for L* is identical to the primary tree for L. The infinite branch indicates that L* cannot be assessed by the c_L-schema, and L* is a singularity of 'true$_{cL}$' – just as L is.[84]

Gauker and Field focus largely on versions of the *repetition* discourse. *Rehabilitation* provides another variant of the strengthened liar.[85] As we saw in Chapter 2, we re-evaluate L itself by an explicitly reflective schema according to which L is true. And as we saw in Chapter 3, if I say in some neutral context that the sentence on the board in 213 is true, my use of 'true' will have L in its scope: my neutral context is non-explicitly reflective with respect to L. In each of these reflective contexts, the token L is evaluated as true. Yet earlier, L is evaluated as not true – that is, not true$_{cL}$. So, as we saw in Chapter 2, 'true' is assessment-sensitive – the semantic value of the token utterance L depends on the context in which it is evaluated.[86]

We can consider more closely a neutral evaluation of L. Suppose Nancy says

(N) The sentence written on the board in room 213 is true,

where the context c_N of N is neutral with respect to L. Nancy's context of utterance is non-explicitly reflective with respect to L, and L is true$_{cN}$, as Nancy says. Yet when we assess N, we might have the intuition that Nancy's utterance is false (that is, false as assessed from our context of evaluation) because L is pathological and so not true. What explains this intuition? We stop our assessment of L too soon: we don't go on to reflect on L, and establish a final value for it. We stop with the assessment of L as pathological, and so not true$_{cL}$. And then we take Nancy to be saying that L is true$_{cL}$, which is indeed false (from our context, or from any context for that matter). But Nancy's associated schema is the c_N-schema, and she is saying that L is true$_{cN}$. The intuition here is the result of going only half-way.

The intuition can be captured more formally in the following way. Suppose we evaluate N by the following procedure: consider the *primary tree* of each of the members of N's determination set. If the tree indicates that the member is pathological, treat that as its final status, and evaluate N accordingly. So L is treated as pathological, so what you say is false. This procedure generalizes in an obvious way. If, for example, you say "Everything Joe says is true", then, according to the present procedure, we consider the primary trees of each sentence Joe says, and if any of the sentences is pathological (but not pathologically tangled with what you say), we take that to be its final status. So if any of Joe's sentences are pathological, what you say is false. From the point of view of the singularity theory, this procedure falls short: we can reason through pathology to a reflectively established value. If Joe produced the Liar sentence L on the board, which is true in your (non-explicitly) reflective context, and everything else Joe says is true, then what you say should be evaluated as true. We'll return to this point when we discuss deflationary accounts of truth in Chapter 10.

III.2 More liars

The singularity theory covers the notion of falsity along with truth. Consider the simplest version of the Liar, where F is written on the board in room 213:

(F) The sentence on the board in room 213 is false.

Let c_F be F's context, and let '$false_{c.F}$' represent the occurrence of 'false' in F, and any coextensive occurrence. Now a falsity c_F-schema comes into play:

s is $false_{cF}$ if and only if it is not the case that S

where 's' is a name of the sentence that replaces 'S'. Instantiating to F, we obtain:

F is false$_{cF}$ if and only if it is not the case that F is false$_{cF}$.

So F cannot be evaluated by the falsity c_F-schema.

The primary representation of F is <type(F),c_F,c_F> where the third element indicates that F is to be evaluated by the falsity c_F-schema. The primary tree for F is:

<type(F),c_F,c_F>
|
<type(F),c_F,c_F>
.
.
.

This branch is infinite, indicating that F is pathological – the attempt to assess it by the c_F-schema breaks down. So F is a singularity of the occurrence of 'false' in F. Once we remove F from the extension of 'false' in F, we can reflectively evaluate F – as false. This corresponds to strengthened reasoning about F. If F were false$_{cF}$, then it would be assessable by the c_F-schema – since it isn't so assessable, it is not the case that F is false$_{cF}$. From this, we conclude:

(€) F is false.

F is false, but not false$_{cF}$. Represent the occurrence of 'false' in € by false$_{c€}$. The $c_{€}$-schema is an r_F-schema, a schema reflective with respect to F. We reach € via this instance of the $c_{€}$-schema:

F is false$_{c€}$ iff it's not the case that the sentence on the board in Caldwell 213 is false$_{cF}$.

We've established the right-hand-side, and we infer the left-hand-side. The simple Liar sentence L is true on reflection because it says it isn't $true_{cL}$ and it isn't $true_{cL}$; the simple Liar sentence F is false on reflection, because it says it's $false_{cF}$ and it isn't $false_{cF}$.

As I noted in Chapter 3, it is easy to construct Liar loops and chains. For example, Jane may say

(J) What Kate is saying now is true.

while Kate is saying

(K) What Jane is saying now is false.

The primary trees for J and K both have infinite branches, and J is identified as a singularity of 'false' in K, and K as a singularity of 'true' in J. A formal treatment of loops and chains is provided in the next chapter.

Another version of the liar is the heterological paradox, generated by the predicate 'true of'. Suppose that I write on the board just this predicate:

(H) predicate on the board not true of itself.

We find that H is pathological – and if H is pathological, then H is a predicate on the board not true of itself. But in the previous sentence, we have produced a predicate token - call it H* - of the same type as H. H* is a *repetition* of H. And H* is not pathological – it has a definite extension, with sole member H. The contextual analysis follows the usual pattern: H* is evaluated by an r_H-schema, a schema reflective with respect to H. We *rehabilitate* H when we evaluate H itself by an r_H-schema, and *iteration* is explained by the oscillation between the two sides of this biconditional.

We will return to more liar-like paradoxes in Chapter 7, where we see how the singularity theory handles the Truth-Teller, Curry's paradox, further liar loops, and paradoxes that do not exhibit circularity.

Endnotes to Chapter 5

---

1. In his discussion of Richard's paradox, Poincaré, unlike Richard, noticed the need for this procedure:

> "In order to classify the integers, or the points in space, I shall consider the sentence which defines each integer or each point. Since it can happen that the same number or the same point can be defined by many sentences, I shall arrange these sentences in alphabetical order and I shall choose the first among these." (Poincaré 1909, p.48)

2. See Richard 1905, in van Heijenoort, pp.142-144.

3. Zermelo's well-ordering theorem states that every set can be well-ordered (see Zermelo 1904 and 1908). Notice that König's argument, if correct, would lead to the conclusion that no non-denumerable set can be well-ordered. König himself recognized the difficulty, pointing out that Cantor's second number class is non-denumerable but can be well-ordered. According to König we should distinguish between *sets* (such as the set of reals), whose members can be well-distinguished, and *classes* (such as Cantor's second number class), whose members cannot be well-distinguished.

4. Russell presents the paradox in Russell 1908 (see van Heijenoort 1967, p.153), and attributes it to G.G. Berry.

5. In his 1905 paper, Richard argues that the phrase has no meaning at the place it occupies, because it refers to a sequence, namely E, that has not yet been defined. In his 1907 paper, Richard gives a different reason: the phrase E is a contradiction-producing diagonal definition.

6. Richard 1905, in van Heijenoort 1967, p.143.

7. Peano 1906, p.357.

8. For an extended discussion of the earliest responses to Richard's paradox, by Richard (1905, 1907), Poincaré (1906, 1909), and Peano (1906), see Simmons 1994a.

9. One author who has suggested that context-sensitivity is relevant to the definability paradoxes is James French (1988). Consider von Neumann's way of identifying the natural numbers with sets: $0=\emptyset$, $1=\{\emptyset\}$, $2=\{\emptyset, \{\emptyset\}\}$, $3=\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, … . Then we can describe the numbers by the following sequence of phrases:

> the empty set,
>
> the set that contains each and every number (and only those numbers) described in the preceding steps,

the set that contains each and every number (and only those numbers) described in the preceding steps,

… .

The endlessly repeating phrase “the set that contains each and every number (and only those numbers) described in the preceding steps” contains less than twenty words, and it describes every natural number greater than zero. And now, French suggests, we have a response to a version of the Berry paradox generated by the phrase “the least natural number not described in fewer than twenty words”. The response is this: every natural number is described in fewer than twenty words, and so the Berry phrase does not name a number.

Clearly, the context-sensitive term “preceding” is crucial to French’s resolution, and he goes on to consider a version of the Berry that eliminates overtly context-sensitive terms - consider for example the Berry phrase “the least natural number which cannot be described in fewer than twenty non-context-dependent words of the English language”. French argues that we need not accept that that this phrase is paradoxical, given the possibility that it contains a word that is context-sensitive. He suggests that “described” is the context-sensitive word, but does little to explain how context shifts arise – he talks vaguely of moves from lower orders of abstraction to higher. Though insufficiently motivated, French’s line is of course broadly congenial to my contextual approach.

10. We could represent R* by the triple $\langle type(R^*), c_{R^*}, c_{R^*} \rangle$. But as we saw in Chapter 4, we have a more perspicuous way of representing repetitions.

11. Richard shows some sensitivity to this iteration in the final paragraph of his original letter:

> “We can make a further remark. The set containing [the elements of] the set E and the number N represents a new set. This new set is denumerably infinite. The number N can be inserted into the set E at a certain rank k if we increase by 1 the rank of each number of rank [equal to or] greater than k. Let us still denote by E the thus modified set. Then the collection of words G will define a number N' *distinct from* N, since the number N now occupies rank k and the digit in the kth decimal place of N' is not equal to the digit in the kth decimal place of the kth number of the set E.” (Richard 1905, in van Heijenoort 1967, pp.143-144. Material within square brackets are editorial interpolations introduced by van Heijenoort to avoid misunderstandings.)

12. Priest has suggested that a contextual approach to the Berry paradox won’t avoid the paradox, because one can explicitly fix the context as one presents the paradox: “There is only a finite number of names with less than 100 words. A fortiori, the number of numbers that I can refer to in this context, c, is finite. Consider the least number that I cannot refer to (in this context). By construction, I cannot refer to it (in c). But I have just referred to it by ‘the least number I cannot refer to in this context’” (Priest 2004, p.119, fn 13).

Utilizing subscripts in the obvious way, Priest’s penultimate sentence here says that I cannot refer$_c$ to the Berry number, i.e. the least number I cannot refer$_c$ to. If the use of ‘referred’ in Priest's final sentence is tied to the context c, as Priest presumably intends, then the final sentence says that I have just referred_c to the Berry number. But this sentence is false, because the Berry phrase ('the least number I cannot refer_c to') has no reference_c conditions, just as C has no denotation_{cC} conditions. The conclusion here should be that the Berry phrase is pathological – we do not reach the contradiction that the Berry phrase both refers_c and does not refer_c, as Priest seems to indicate. So the paradox is not reinstated, even though the context is explicitly fixed. What is established is the pathology of the Berry phrase, and according to the contextual approach, we can go on to reason past this pathology.

The contextual approach locates a grain of truth in what Priest says in the final sentence. If we suppose that all other referring expressions with fewer than 100 words that I can produce in context c are unproblematic, then my use of the Berry phrase can be evaluated by a *reflective* schema which takes into account its pathology – and then the Berry phrase will indeed refer to a number. Compare: when I produce C, in the context c_C, I have produced an expression that denotes_{cE} $\pi+6$ – C does denote when evaluated by a reflective schema. In parallel, in the Berry case I produce, in context c, a phrase that does refer to the least number I cannot refer_c to – as long as the phrase is reflectively evaluated.

13. For Zermelo's characterization of *definite,* see Zermelo 1908b, in van Heijenoort 1967, p.202. For Skolem's clarification, see Skolem 1922, p.292).

14. Zermelo 1908b, in van Heijenoort, p.201.

15. *ibid*.

16. *ibid*.

17. The iterative conception was first presented in Zermelo 1930, and subsequently in Boolos 1971 and Scott 1974.

18. For a detailed discussion of the iterative conception of set, see Boolos 1971.

19. Since Separation "furnishes a substitute" (Zermelo 1908b, in van Heijenoort 1967, p.200) for Naive Comprehension, it might appear that ZF maintains a restricted version of the predicative conception of set, according to which a set is an extension of a predicate. But the appearance is misleading. Separation is a residue of the predicative conception, but it provides only a limited description of the intended universe. The sets of the ZF hierarchy outrun the predicates of our language, and consequently ZF does not embody the predicative conception. Frank Drake writes:

> "This axiom can be regarded as an attempt to say that we intend, at each stage of the cumulative structure, to take *every* collection whose members have already been formed as a set at the next level. But we have only the formulas of our language with which to describe the collections, and this limits the effect of this axiom." (Drake 1974, p.9. Maddy cites Drake in Maddy 1983, p.121.)

20. The 'lasso' figure can be found in Boolos 1971, where it is attributed to Kripke.

21. ZF must meet the challenge that it is no more than an *ad hoc* repair to an inconsistent theory, and so unsuited to play any foundational role. And given the Skolem-Lowenheim theorems, we cannot take the notion of set to be implicitly defined by the axioms of ZF: the axioms do not determine a unique universe of sets, and there is an unavoidable relativity in basic notions such as *enumerability*, *finite*, *infinite*, *cardinality*, and *power set*. But the iterative conception shows that there *is* a unifying picture of the universe of sets, and the axioms of ZFC are an articulation of that universe. For more discussion, see Giaquinto 2002, Part VI.

22. Frege 1903, Appendix.

23. Zermelo 1908b, p.201. Objects in B are either sets or urelements (objects which are not sets, but which can be members of sets).

24. *op. cit.*, p.203.

25. Skolem 1922, p292.

26. Except for the axiom of infinity, these are the axioms of Zermelo 1908b, together with the axiom of replacement and the axiom of foundation. According to Zermelo, the axiom of infinity does not belong to 'general' set theory.

27. According to Zermelo, the paradoxes

> "are only apparent 'contradictions', and depend solely on confusing *set theory itself*, which is not categorically determined by its axioms, with individual *models* representing it. What appears as an 'ultrafinite non- or super-set' in one model is, in the succeeding model, a perfectly good, valid set with both a cardinal number and an ordinal type, and is itself a foundation stone for the construction of a new domain." *op. cit.*, p.1233.

27. *op. cit.*, p.1233.

29. As noted in Chapter 1 (note 8), Giaquinto takes Zermelo's relativistic line about sets and classes to provide a defense of the ZF resolution of Russell's paradox -- see Giaquinto 2002, pp.214-8.

30. Hallett 1996, p.1212. Hallett also points out that there is little plausibility in the claim that B is a set because all its elements (other than the urelemente) are sets (*ibid*.). There is no 'self-reproductive process' here, unlike the case of the ordinals, where whenever we have 'all' the ordinals, this very collection gives rise to a new ordinal.

31. Zermelo recognizes the special character of the sequence of boundary numbers: "… the *existence of an unbounded sequence of boundary numbers* must be postulated as a *new axiom* of 'meta-set theory'." (Zermelo 1930, p.1233**)**

32. D.A. Martin, circulated photocopy, [n.d.]. Martin is responding to Parsons 1974, in which the set/class distinction is treated relativistically.

33. The axioms of Quine's NF are the axiom of Extensionality and an axiom schema of Naive Comprehension restricted to *stratified* instances: that is, any occurrence of '∈' must be flanked by variables with consecutive ascending indices. Since the formula 'x=x' is stratified (vacuously so, because '∈' does not appear in it), we *can* prove the existence of a universal set V directly from the comprehension schema.

34. In NF we cannot prove from the comprehension schema that there is a set of the non-self-membered sets, because 'x∉x' is unstratified.

35. The paradox runs as follows. Suppose there is a set *WF* of exactly the well-founded sets. Is *WF* well-founded? Suppose it is; then it is a member of itself. But then *WF* is the first link in an infinite membership chain (...∈*WF*∈*WF*∈*WF*), and so is not well-founded. So suppose on the other hand that *WF* is not well-founded. Then it is the first link on an endless membership chain. But then at least one of its members must also be the first link on an endless membership chain. That is, one of its members is not well-founded - but *WF* is the set of exactly the well-founded sets. Either way we get a contradiction, and we are landed in paradox. The paradox is avoided in NF, since the well-founded sets do not form a set. (Notice that in NF, the universal set V is not the set of exactly the well-founded sets, for NF admits sets that are not well-founded. V is one of them, since V is a self-member.)

Given various substantial difficulties with NF – for example, there are sets, such as V, for which Cantor's theorem does not hold, the relation of *less-to-greater* among cardinal numbers is not a well-ordering, and the Axiom of Choice fails – Quine abandoned NF in favor of ML. ML distinguishes sets and ultimate classes, and there is no universal set.

36. In the systems of Church 1974 and Mitchell 1976 there is, for example, no set of the well-founded sets, no set of the non-self-membered sets, and no set of all ordinals.

37. See Vopenka and Hajek 1972, and Vopenka 1979. Vopenka and Hayek developed their theory of semisets for a set theory without a universal set.

38. Vopenka's leading examples of semisets turn on the phenomenon of vagueness. The class of all apes, the class of all living men, and the class of bald men are proper semisets, because the applications of the predicates 'ape', 'living man' and 'bald man' do not have crisp boundaries (see Vopenka 1979, pp.33-34).

39. There are other examples in NF of subclasses of sets that are not sets, including some that are finite (see Forster 1992, pp.30-1).

40. As we saw in note 37, the extensions of 'ape', 'living man' and 'bald man' are examples given by Vopenka of semisets. Vopenka writes: "… we meet proper semisets whenever in considering a property of some objects we emphasize its intension rather than its extension." Vopenka 1979, p.34.

41. von Neumann 1925, p.401.

42. Gödel 1940 also treats sets as certain classes; in contrast, Bernays 1937 and 1941 keep sets and classes separate.

43. *op. cit*. p.401, p.403. Instead of working directly in terms of sets, von Neumann couches his 1925 paper in terms of functions and arguments. So rather than saying that sets cannot be members, von Neumann says that proper classes -- are "incapable of being arguments". Bernays (1937, 1941) recast von Neumann's system in more familiar set-theoretical terms. And in Skolem 1938, von Neumann's axioms are presented in the first-order predicate calculus (see pp.32-4).

44. It might seem natural to restrict φ to formulas that contain no class variables, only set variables. However this restriction would make the system cumbersome. Levy *et al* illustrate the point as follows (see Müller 1976, pp.180-1). Consider the statement:

(+) For every class A there exists a class B of all sets that are not members of A.

Now (+) is true, because whatever condition φ determines the class A, B is determined by '-φ(x)'. But without class variables, we cannot prove (+) in general (though we can prove every instance of (+)). With class variables we can prove (+). Given any class A, we can put '$x \notin A$' for φ in the Axiom of Comprehension for classes, and thereby demonstrate the existence of the complement class B.

45. And there are technical costs too. It is provable that in von Neumann's system one cannot prove all instances of the induction schema – see Mostowski 1951.

46. See Levy *et al*, in Müller 1976, p.201, and Quine 1963, p.51.

47. See van Heijenoort 1967, pp.399-400.

48. Wang 1949, Kelley 1955, Morse 1965. Quine suggests such an axiom schema as part of his system ML (see Quine 1951).

49. See Müller 1976, p.201.

49. This system is consistent if $ZF^{\#}$ is consistent (see below for $ZF^{\#}$).

51. Levy *et al*, in Müller 1976, p.202.

52. Levy and Vaught have shown that these power classes also exist in Ackermann's system of classes (see Levy 1959, Levy-Vaught 1961; for Ackermann's system, see Ackermann 1956).

53. See Levy *et al*, in Müller 1976, pp.201ff.

54. See Aczel 1987. Aczel's set theory is used by Barwise and Etchemendy to model circular propositions that produce Liar paradoxes (see Barwise and Etchemendy 1987).

55. Frege 1979, p.181 (‘On Schoenflies: Die Logischen Paradoxien der Mengenlehre’). In a similar vein, Frege writes:

> “I myself was long reluctant to recognize ranges of values and hence classes; but I saw no other possibility of placing arithmetic on a logical foundation. But the question is, How do we apprehend logical objects? And I have found no other answer than this, We apprehend them as extensions of concepts...” (Frege 1980, pp. 140-1.)

In the course of responding to Russell's paradox, Frege writes:

> "And even now I do not see how arithmetic can be scientifically established; how numbers can be apprehended as logical objects, and brought under review; unless we are permitted - at least conditionally - to pass from a concept to its extension." (Appendix to vol. ii of *Grundgesetze*, in Geach and Black 1952, p.234.)

56. Geach and Black 1952, p.106. Frege’s criticizes Schröder for holding the view that "classes consist of single things, are collections of individuals", and goes on:

> "Only because classes are determined by the properties that individuals in them are to have, and because we use phrases like this: 'the class of objects that are b'; only so does it become possible to express thoughts in general by stating relations between classes; only so do we get a logic." (*op*. *cit*., pp.104-5)

57. See, for example *op*. *cit*., pp.149-150 (from the Introduction to *Grundgesetze*), and *op*. *cit*., p.102.

58. Frege writes:

> “He [Grassmann] forms classes or concepts by logical addition. He would e.g. define 'continent' as 'Europe or Asia [or Africa] or America or Australia'. But it is surely a highly arbitrary procedure to form concepts merely by assembling individuals, and one devoid of significance for actual thinking unless the objects are held together by having characteristics in common. It is precisely these which constitute the essence of the concept. Indeed one can form concepts under which no object falls, where it might perhaps require lengthy investigation to discover that this was so. Moreover a concept, such as that of number, can apply to infinitely many individuals. Such a concept would never be attained by logical addition.” (Frege 1979, p.34)

59. Russell 1919, p.12.

60. Forster 1992, p.1.

61. *op*. *cit*., p.11 (the emphases are Forster's).

62. This question is also taken up in Parsons 1974b.

63. We noted above that Aczel's set theory includes an anti-Foundation axiom. But we saw that Aczel's theory cannot serve as a theory of extensions, because it does not provide extensions for the predicates 'set', 'non-self-membered set', 'well-founded set', etc, any more than ZF does.

64. Zermelo 1904, p.516.

65. For a view typical of the French constructivists, see Lebesgue 1905, in Moore 1982, p.314 and pp.316-7. See also Russell 1911, translated in Grattan-Guinness 1977, pp.161-74.

66. So the iterative conception provides a straightforward justification of the Axiom of Choice. It is, then, a drawback of Boolos's treatment of the iterative conception that it turns out to be neutral with respect to the Axiom of Choice. Boolos develops an intuitive 'stage theory', which, it is claimed, embodies the iterative conception - and from which many of the axioms of ZF follow. But the Axiom of Choice itself does not follow from the stage theory, and Boolos concludes that "the justification for its acceptance is not to be found in the iterative conception of set" (Boolos 1971, in Putnam and Benacerraf 1983, p.502.)

67. Hadamard *et al* write that the notion of a describable mapping or correspondence "belongs to the field of psychology and concerns a property of our minds" - it is "outside mathematics" (Hadamard *et al* 1905/1982, in Moore 1982, p.312.) Hadamard *et al* had no doubt that the choice mappings and choice sets exist. And in opposition to Russell, Ramsey wrote: "The possibility of indefinable classes and relations in extension is an essential part of the extensional attitude of modern mathematics ..., and that it is neglected in *Principia Mathematica* is the first of the three great defects in that work." (Ramsey 1925, p23).

68. See, for example, the presentation of a version of von Neumann's set theory, in Levy *et al*, in Müller 1976, p.178. Maddy points out that Bernays and Gödel took von Neumann's proper classes as substitutes for Zermelo's definite conditions (Maddy 1983, p.121; see Bernays 1937, Gödel 1940). And Levy *et al* point out that Quine and Bernays were motivated by the idea of replacing the metamathematical notion of condition by the mathematical notion of class (Levy *et al*, in Müller 1976, p.196; see Quine 1963, Bernays 1958.)

69. von Neumann 1925, in van Heijenoort 1967, p.404.

70. Ackermann's system is in a strong sense equivalent to ZF - see Levy *et al*, in Müller 1976, pp.210-2.

71. Contextual approaches to truth motivated by the strengthened liar can be found in Parsons (1974), Burge (1979), Simmons (1993), Glanzberg (2001). Critical discussion can be found in Gauker (2006) and Field (2008, Chapter 14).

72. Gauker 2006, pp.395-396. My presentation of the strengthened liar differs from Gauker's in the following respects. Gauker's presentation includes (a terminological variant of) the identity

L = 'L is not true'

But I've avoided this identity, since we are dealing with tokens (or sentence types in a context).

Notice that if we say L = ‘L is not true’, then it seems we should also say (2) = ‘L is not true’. But of course it is false that L= (2), since they are different tokens. What we can say is that L and (2) say the same thing, or have the same content.

Gauker also uses quote names in moving, by Semantic Ascent, from

(2) L is not true

to

(3) ‘L is not true’ is true.

Again, I avoid quote names in my presentation. (2) is a token of the type ‘L is not true’ -- but the quote name ‘L is not true’ is not a name of the token (2), any more than it’s a name for the token L. In my presentation, the move by Semantic Ascent is from (2) to

(3) (2) is true.

Here ‘(2)’ is a legitimate name, and Semantic Ascent takes us from the right-hand-side of the biconditional *(2) is true iff L is not true* to the left-hand-side.

73. *ibid.*

74. *op. cit.*, p.397.

75. *op. cit.*, pp.397-8.

76. Gauker expresses doubt that L says anything, since it fails to have truth_{cL} conditions (*op. cit.*, p.400). But it does say something – it says it’s not true_{cL}. It cannot be assessed by the c_L-schema, but from that it does not follow that it fails to say anything. It says that it’s not in the extension of ‘true_{cL}’, and indeed it cannot be, on pain of contradiction. So it doesn’t just say something, it says something true (the same true thing that L* says). But the evaluation of L as true requires a suitably reflective schema.

77. *op. cit.*, p.400.

78. With regard to a reflective schema like the c_Σ-schema, Gauker raises two questions: “… with what right may we take this biconditional for granted? … And second, with what right may we take for granted the right-hand side?” (op. cit., p.401). On the first question: my project here is to provide an analysis that respects the strengthened reasoning as intuitive and valid. So it’s a matter of figuring out exactly what we’re doing when we carry out this reasoning. According to the contextual analysis, we do as a matter of fact employ the reflective c_Σ-schema in the later stages of the discourse. And this schema, we find, applies to both L and L*, and the argument as analyzed is valid. (That’s not to say that the schema will never break down. The predicate ‘$\text{true}_{c\Sigma}$’ has its own singularities – for example, certain perverse anaphoric additions -- to which the c_Σ-schema fails to apply. But the application of the schema to L and L* succeeds.) As to the second question: again, we establish that L is not true_{cL} in the course of our reasoning, and that *is* the right-hand-side.

79. Field 2008, pp.211-2.

80. *op. cit.*,, p.212.

81. *ibid*.

82. *op*. *cit*., p.213.

83. *ibid*.

84. In Chapter 8, I argue that the phenomenon of repetition poses a problem for Field's own *paracomplete* theory of truth.

85. Burge works with a version of *rehabilitation,* in his 1979.

86. Assessment-sensitivity also figures in Scharp's theory of truth, but in a very different way. Scharp provides two theories: a *prescriptive theory,* according to which truth is an inconsistent concept to be replaced by the concepts ascending truth and descending truth, and a *descriptive theory*, a theory of the inconsistent concept of truth that appeals to ascending truth and descending truth. Scharp's descriptive theory takes the truth values of sentences containing 'true' to be relative to a context of utterance *and* a context of assessment. The descriptive theory gives the *ascending truth* values and the *descending truth* values of sentences containing 'true' (not their *truth* values). And it gives these values at a context of use *from a context of assessment*, because 'true' is treated as assessment-sensitive. According to Scharp's assessment-sensitivity approach (Scharp 2013, p.250ff), given a sentence containing 'true', a context of use, and a context of assessment, the context of assessment determines how to read the occurrence of 'true', either as 'ascending true' or 'descending true' – and then, given that reading, it's a further matter as to whether the sentence is to be assessed for ascending truth or descending truth. (On an alternative approach that Scharp considers -- the non-indexical contextualist option (p.252ff) -- the context of assessment determines both the reading of 'true' and whether it is assessed for ascending truth or descending truth. This approach is simpler but less versatile, and Scharp endorses the assessment-sensitive approach (p.256)). For a full account of the descriptive theory, see Scharp 2013, Chapter 9.

Though assessment-sensitivity is a feature of both Scharp's theory and my contextual account, there are several major differences. First, and most obviously, Scharp's account of assessment-sensitivity is in terms of the notions that replace truth, ascending truth and descending truth, notions not available to the ordinary speaker. On my contextual account, only our ordinary truth predicate is in play, and the assessment-sensitivity of sentences involving 'true' is motivated by natural reasoning about the liar. Second, Scharp's account does not make it clear how the reading of 'true' (as either 'ascending true' or 'descending true') is determined by the context of assessment. Scharp sometimes talks in terms of a speaker 'deciding' or 'choosing' one or the other (see p.251), but, as Scharp observes, one may wonder what motivates one of these choices over the other (*ibid*.) – and certainly the ordinary speaker would not have the conceptual resources to make such a choice. On my account, the shift in the evaluating schema occurs in the course of natural reasoning (*repetition* or *rehabilitation*), and is explained in terms of a contextual parameter, *reflective status*. Third, Scharp's theory embraces a version of semantic relativism (see p.241): a sentence containing 'true' receives a value (ascending true or descending true) only relative to the reading given to 'true' ('ascending true' or 'descending true'). In contrast, on my account, assessment-sensitivity lends no support to relativism, as we saw in Chapter 2.5.