8

Semantical and Logical Paradox

KEITH SIMMONS

1 Introduction

Consider the following array:

	'monosyllabic'	'French'	'inanimate'	'infinite'	. . .
'monosyllabic'	f	t	f	f	. . .
'French'	f	f	f	f	. . .
'inanimate'	t	t	t	t	. . .
'infinite'	f	f	f	f	. . .
⋮	⋮	⋮	⋮	⋮	⋱

Down the side and along the top are the 1-place predicates of English, taken in the same order. In each box, we put t or f according to whether the predicate at the side is true of the predicate at the top. We obtain rows of ts and fs. For example, the row of values associated with 'monosyllabic' is: ftff. . . . Consider the diagonal of values from the top left towards the bottom right: fftf. . . . Observe that each f in this *diagonal sequence* corresponds to a predicate false of itself ('monosyllabic,' 'French' and 'infinite' are each false of themselves). Now form the *antidiagonal sequence* by changing each f in the diagonal sequence to a t, and each t to an f. We obtain the sequence ttft . . . , where now each t corresponds to a predicate false of itself. Notice that this antidiagonal sequence cannot occur as a row: it differs from the first row in the first place, from the second row in the second place, and, in general, from the nth row in the nth place. So there can be no predicate of English true of exactly those English predicates false of themselves – for if there were such a predicate, its associated row would be our antidiagonal sequence. But there *is* such a predicate – consider the predicate we've just used in the previous sentence, namely ‘English predicate false of itself,’ or ‘heterological’ for short. We are landed in paradox.

Now make some changes to the array. Replace each predicate by its extension, and in each box put ‘$\in$’ or ‘$\notin$’ according to whether the extension at the side has for a member the extension at the top. On certain natural assumptions, we obtain this array:

	extension of ‘monosyllabic’	extension of ‘French’	extension of ‘inanimate’	extension of ‘infinite’	. . .
extension of ‘monosyllabic’	$\notin$	$\notin$	$\notin$	$\notin$	. . .
extension of ‘French’	$\notin$	$\notin$	$\notin$	$\notin$	. . .
extension of ‘inanimate’	$\in$	$\in$	$\in$	$\in$	. . .
extension of ‘infinite’	$\notin$	$\in$	$\in$	$\in$	. . .
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

(In particular, we assume there are finitely many monosyllabic words, and infinitely many French objects – consider the totality of French sentences. And we assume there are infinitely many inanimate things, and infinitely many infinite things – just consider the infinitely many infinite, and inanimate, extensions generated by the predicates ‘natural number greater than 1,’ ‘natural number greater than 2,’ and so on.) Again we can form the diagonal sequence $\notin\notin\in\in$. . . . The antidiagonal sequence is $\in\in\notin\notin$ . . . , where each $\in$ corresponds to an extension that is not a member of itself (such as the extension of ‘monosyllabic’ and the extension of ‘French’). Again, this antidiagonal sequence cannot occur as a row. So, on pain of a contradiction, there can be no English predicate whose extension is exactly the *non-self-membered extensions of predicates of English*. But the italicized predicate in the previous predicate *is* such a predicate, and we are landed in paradox again.

Our first paradox – the heterological paradox – is a member of the family of *Liar paradoxes*. The Liar takes many forms: for example, versions of the Liar are generated by the sentences “This sentence is false,” “This sentence is not true,” and “I am lying now.” All forms of the Liar turn on the semantic notions of truth or falsity, and so they in turn are members of the family of *semantic paradoxes*. Other members of this extended family turn on the notion of reference or denotation – these are the so-called ‘definability paradoxes’ due to Richard, König and Berry (see Richard 1905; König 1905; and for the Berry, Russell 1908). Consider, for example, Berry’s paradox. There are only a finite number of English expressions with fewer than 19 syllables, and some of these (like ‘the square of 3’) denote integers. But there are infinitely many integers. Let k be *the least integer not denoted by an English expression in fewer than 19 syllables*. This italicized phrase denotes k, but it has fewer than 19 syllables – and we've reached a contradiction.

Our second paradox is a version of Russell's paradox, couched in terms of extensions. Russell's paradox also arises for sets – consider the set of exactly the non-self-membered sets, and ask whether or not it is a self-member. This version of Russell's paradox is one of several set-theoretical paradoxes discovered at the turn of the twentieth century. Among these are Burali-Forti's paradox, turning on the set of all ordinal numbers, and Cantor's paradox, concerning the universal set, the set of all sets.

Following Ramsey (1925), it has become standard to divide the paradoxes into two groups: the *semantical* paradoxes (such as the Liar and the definability paradoxes), and the *logical* paradoxes (such as Russell's, Burali-Forti's, and Cantor's). And the attempts to resolve the two kinds of paradoxes have tended to go their separate ways. There is something to this division: the semantical paradoxes arise from ordinary terms of English, like 'true' and 'denotes', while the set-theoretical paradoxes arise in the setting of a mathematical language, and turn on technical notions like *set* and *ordinal number*.

Nevertheless, we should be wary of the division, at least the way Ramsey draws it. As we have seen, the heterological paradox and Russell's paradox have a shared structure: each is generated by a *diagonal argument*. Diagonal arguments establish positive theorems – for example, Gödel's first incompleteness theorem, Tarski's indefinability theorem, and many theorems of recursion theory. But they also generate paradoxes. (For more on the diagonal argument, see Simmons 1993.) The shared diagonal structure of the heterological paradox and Russell's paradox encourages the search for a common resolution. Moreover, Russell's paradox for extensions is tied to predication – and that encourages the thought that it belongs in the category of semantical paradox, along with the heterological paradox.

So there may be a question about how best to classify the paradoxes. But there is no doubt about their tremendous significance: they have forced logicians and philosophers to rework the foundations of semantics and set theory.

2 Semantic Paradoxes: Some Proposals

The hierarchy

Think back to our first paradoxical array, where the top and the side were composed by *all* the 1-place predicates of English, *including* the problematic 'English predicate false of itself.' We can escape paradox if we restrict the array in some suitable way, so that this and other paradox-producing predicates are excluded. Suppose the restricted side and top is the collection D of semantically unproblematic 1-place predicates of English. In particular, the predicate 'English predicate in D false of itself' is excluded from D, on pain of paradox. Here is an application of Russell's Vicious Circle Principle: "Whatever involves *all* of a collection must not be one of the collection" (Russell 1908: 155). We might think of the predicate 'English predicate in D false of itself' as standing above the collection of predicates that it involves.

These ideas may lead us to a *hierarchical* account of truth and falsity. One such account runs as follows. At the first level of the hierarchy are the expressions of English that do not contain any semantic terms (predicates like 'monosyllabic' and sentences like 'Aardvarks amble'). At the second level we find these first-level expressions together with semantical predicates, like 'true_1,' 'false_1,' 'true_1 of itself,' and 'false_1 of itself,' *which apply only to sentences and predicates of the first level*. We can think of the second-level language as a *metalanguage* for the object language of the first level – the metalanguage contains the semantical terms that apply to the object language. At the third level we find all the second-level expressions (including all the first-level expressions), together with semantical predicates, like 'true_2,' 'false_2,' 'true_2 of itself,' and 'false_2 of itself,' *which apply only to sentences and predicates of the second level*. And so on.

Now semantical paradox does not arise. For example, we can no longer generate the heterological paradox. There is no *absolute* predicate 'English predicate false of itself,' but rather *relativized* predicates of the form 'English predicate false_α of itself' for some ordinal α. This predicate is of level $\alpha + 1$, and applies only to predicates of level α. So it does not apply to itself – and a contradiction is no longer forthcoming.

Or consider the Liar sentence:

(L) (L) is not true.

Here we generate a contradiction by observing that

(1) "(L) is not true" is true if and only if (L) is not true.

This is an instance of Tarski's famous truth-schema:

(T) X is true if and only if p,

where 'p' abbreviates a sentence, and 'X' is a name of that sentence (see Tarski 1944: 15). Given (1), and given that '(L) is not true' just is the sentence (L), we may infer:

(2) (L) is true if and only if (L) is not true,

from which a contradiction immediately follows.

But if we adopt the hierarchical view, this derivation is blocked. Just as the truth (and falsity) predicates are always relativized to a level, so is the truth-schema. The occurrences of 'true' in (L) and in (T) are relativized to some level. So (L) is to be understood as '(L) is not true_β,' for some ordinal β. The T-schema associated with 'true_β' is:

(T_β) X is true_β if and only if p,

for some ordinal β, where 'p' abbreviates a sentence of level β, and 'X' names that sentence. Observe that, according to the hierarchical line, (L) is a sentence of level $\beta + 1$ and not of level β. So it may not be substituted for p in the schema (T_β), and this blocks the Liar reasoning.

How attractive is the hierarchical resolution of semantical paradox? It faces a number of serious difficulties. First, the splitting of 'true' and 'false' into an infinity of distinct, stratified predicates seems to go against the spirit of a natural language like English. English doesn't seem to be stratified, and the predicate 'true' appears to be univocal. Before his eventual endorsement of the hierarchical approach, Russell himself described it as "harsh and highly artificial" (Russell 1903: 528).

Second, the stratification of 'true' (and 'false') involves massive restrictions on occurrences of 'true.' On a standard hierarchical line, Tim's utterance of "Aardvarks amble" is of level 1; Joanne's utterance of " 'Aardvarks amble' is true" is of level 2; and so on, through the levels. Joanne's use of 'true' has in its extension all sentences of level 1 and no others. So all sentences of level 2 and beyond are excluded from the extension of Joanne's use of 'true' (and any use of 'true' in a sentence of level 2). Gödel remarked of Russell's type theory that "each concept is significant only . . . for an infinitely small portion of objects" (Gödel 1944: 149). A similar point can be made here about the hierarchical line: an ordinary use of 'true' will apply to only a fraction of all the truths.

Third, the hierarchical resolution invites a *revenge Liar* – a version of semantical paradox couched in the very terms of the resolution itself. Consider the sentence:

(3) This sentence is not true at any level of the hierarchy.

Suppose (3) is true – that is, on the hierarchical line, true at some level ρ. Then what (3) says is the case, and so (3) is not true at any level, and, in particular, not true at level ρ. Suppose, on the other hand, that (3) is not true at any level. But that is just what (3) says – so (3) is true (at some level). We obtain a contradiction either way: we have traded the old paradoxes for a new one.

Truth-value gaps

Given these difficulties, we might wonder if we can dispense with the hierarchy. In the terms of our first array, let us admit 'English predicate false of itself,' or 'heterological,' to the side and top, and make adjustments elsewhere. Now we should ask: what value can we put in the 'heterological'/'heterological' box in the leading diagonal? On pain of contradiction we cannot put 't' or 'f' in this box. So a natural thought is to appeal to *truth-value gaps*, and say that the predicate 'heterological' is neither true nor false of itself. And we can put 'u,' say, in the 'heterological'/'heterological' box. Now suppose we form the antidiagonal by converting each t to an f, each f to a t, and leaving each u unchanged. Observe that the antidiagonal is identical to the row associated with 'heterological,' and no contradiction arises. Contradiction arises if we assume heterological is true or false of itself – but if it is neither, we escape the paradox. Similarly for Liar sentences; for example, the sentence 'This sentence is false' only generates a contradiction if we assume it is either true or false.

The claim that Liar sentences are gappy seems natural enough – after all, the assumption that they are true or false leads to a contradiction. Moreover, one can motivate gaps independently of the Liar (e.g. by appeal to presupposition theory, or category considerations, or vagueness).

With gaps on board, we can allow the predicate 'English predicate false of itself' to belong to the collection of English predicates – we have no need to invoke Russell's

Vicious Circle Principle. More generally, Kripke (1975) has shown that, if we admit truth-value gaps, it is possible for a language to contain its own truth-predicate. By a fixed-point construction, Kripke obtains a language – call it L_σ – that contains the predicate 'true-in-L_σ,' the extension of which is exactly the true sentences of L_σ. And similarly for the predicate 'false-in-L_σ.' The language L_σ exhibits a striking degree of semantic closure. L_σ has the capacity to express its own concepts of truth and falsity; there is no need to ascend to a metalanguage.

So truth-value gaps are natural enough, and it might appear that they allow us to dispense with the hierarchy. But a moment's reflection shows that any such appearance is deceptive. The 'truth-value gap' approach to semantic paradox faces its own revenge Liar, couched in terms of gaps. Consider the sentence:

(4) This sentence is either false or gappy.

This Liar sentence generates a contradiction whether we assume it is true, false, *or gappy*. In particular, if (4) is gappy, then it is either false or gappy – but that's what (4) says, so it's true. A similar paradox is produced by the sentence:

(5) This sentence is not true,

as long as 'not true' is taken in a suitably wide sense, as coextensive with 'false or gappy' (and not with 'false'). This is a perfectly natural sense of 'not true.' False sentences are not true, of course, but so are gappy sentences – indeed, gappy sentences are, by definition, not true (and not false).

There is a revenge version of the heterological paradox too – just consider the predicate 'English predicate false or neither true nor false of itself' ('superheterological' for short). Or to put it in terms of our array: form the antidiagonal by changing each t to an f, each f to a t, and each u to a t. Now this antidiagonal cannot occur as a row of the array, on pain of contradiction. And yet this antidiagonal sequence just is the row associated with 'superheterological.'

Perhaps we must appeal to the Vicious Circle Principle again, and exclude 'superheterological' from the class of English predicates that it involves. And that would lead us back to the hierarchy. Similarly with Kripke's language L_σ. Although L_σ contains its own truth and falsity predicates, it does not contain 'neither true-in-L_σ nor false-in-L_σ,' or 'not true in L_σ' (in the appropriately wide sense). If we admit these predicates into L_σ, the revenge Liar returns. According to the truth-gap approach, Liar sentences are gappy, and they are not true; however, we cannot say so in L_σ, but only in a semantically richer metalanguage. The language in which we state the gap account, in which we express the notion of a truth-value gap, must be regarded as a metalanguage for L_σ (see Kripke 1975: 79–80, and fn. 34).

Return of the hierarchy?

The point here can be generalized. Suppose I offer a resolution of semantical paradox that makes no appeal to a hierarchy. Let $\mathcal{L}$ be the object language, the language that my semantical theory is a theory of. And let $\mathcal{L}_T$ be the language in which I state my theory. We can ask: is $\mathcal{L}_T$ a metalanguage for $\mathcal{L}$, on pain of semantical paradox? This is a crucial question, for if the answer is affirmative, then I have not dispensed with the hierarchy, and I have not dealt with semantical paradox in all its forms.

It is a question we can raise not only for Kripke's theory, but for a wide variety of non-hierarchical theories of truth, such as the revision theory (see Gupta 1982; Gupta and Belnap 1993; Herzberger 1982), McGee's treatment of 'true' as a vague predicate (McGee 1990), and Feferman's type-free theory of truth (Feferman 1982). For example, a key notion of the revision theory is that of *stable truth*. The leading idea is that Liar sentences are unstable: if we ascribe truth to the Liar sentence (L), we must revise that ascription, declaring (L) untrue, and then in turn revise that ascription, declaring (L) true, and so on indefinitely. We can ask whether the notion of stable truth must be confined to a metalanguage, on pain of the revenge Liar generated by

(S) (S) is not stably true.

Parallel questions can be raised for McGee's notion of *definite truth*, and for the notion of 'not true' in Feferman's theory (where negation is classical). Observe that all these notions at issue – truth-value gaps, stable truth, definite truth, untruth – are natural enough, and so it is all the more urgent that a purported solution to the Liar come to grips with them. (For an extended discussion of these matters, see Simmons 1993.)

Dialetheism

We have seen that there are serious difficulties with the hierarchical approach. Now suppose we become convinced that nonhierarchical approaches cannot really avoid the hierarchy. In the face of this dilemma, we might seek more radical measures. According to *dialetheism*, Liar sentences are *both true and false* (see, e.g., Priest 1979, 1984). According to Priest, once we admit such truth-value 'gluts', we may dispense with the object language/metalanguage distinction altogether (see Priest 1984: 161). Of course, dialetheism requires that we abandon classical principles of semantics and logic – but only for a certain class of pathological cases, like the Liar family. Perhaps we can cordon off the paradoxical sentences, so that truth-value gluts will be the exception rather than the rule, and classical principles will hold good everywhere else.

But it may not be so clear that the dialetheist can prevent the spread of pathology. One dialetheist account of the truth conditions of 'A is true' and 'A is false' is summed up by these tables:

A	*A is true*
t	t
p	p
f	f

A	*A is false*
t	f
p	p
f	t

where 'p' abbreviates 'paradoxical' (i.e. 'true and false') (see Priest 1979). Let L be a Liar sentence. Then L is both true and false. By the truth tables,

L is true $\leftrightarrow$ L
and L is false $\leftrightarrow$ L.

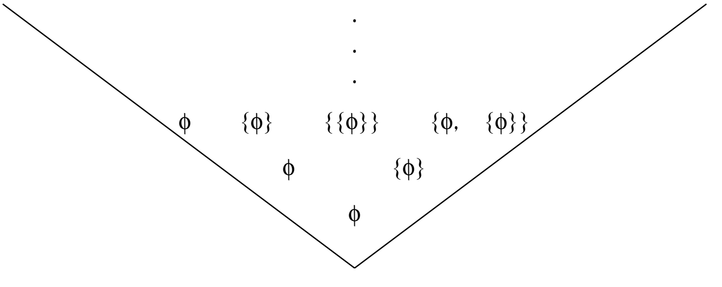
So 'L is true' and 'L is false' are paradoxical. Since, according to the present dialetheist account, the conjunction of two paradoxical sentences is paradoxical, 'L is true and L is false' – that is, 'L is paradoxical' – is paradoxical. So a defining claim of the dialetheist account, that L is paradoxical, is itself paradoxical. The theory itself is not immune from paradoxical assertions. And perhaps this should give us pause.

3 Sets and Extensions

Recall Russell's paradox for sets. Given the set R of exactly the non-self-membered sets, we obtain a contradiction if we assume it is self-membered, and if we assume it isn't. Nowadays, this paradox is no longer considered a real threat: it does not arise in the received set theory, Zermelo-Fraenkel set theory (ZF).

ZF set theory embodies the *combinatorial* or *iterative* conception of set (see Boolos 1971). 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 ϕ. Every pure set appears somewhere in this endless cumulative hierarchy:



Observe that no ZF set is a self-member. So if the Russell set R existed it would be the universal set. But there is no universal set, since there is no end to the hierarchy. In this way, Cantor's paradox is avoided. And since it follows that there is no set R, Russell's paradox for sets is also avoided. (Similarly with Burali-Forti's paradox: there is no set of *all* ordinal numbers.)

ZF does provide a consistent set-theoretical basis for mathematics. But there are costs. 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 since there is no universal set, ZF does not provide an extension for '$x = x$.' 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. Note further that in ZF we quantify over sets, and so we need a domain of quantification; but again no set in the hierarchy can serve as this domain.

Such considerations have led some to explore the prospects of a set theory with a universal set (see, e.g., Quine 1937). Thus far those prospects do not seem very bright, at least if we are after a set theory that is plausible and intuitive. A more entrenched response has been to introduce another kind of collection: *classes* or *proper classes*. Proper classes are collections 'too big' to be sets; there is, for example, a proper class of all sets and a proper class of all ordinals. (Proper classes were first explicitly introduced in von Neumann 1925; for a recent discussion, see Maddy 1983.) Of course, a version of Russell's paradox threatens proper classes too. Von Neumann's way out placed a restriction on proper classes: they cannot themselves *be* members. This restriction is severe – we cannot even form the unit class of a proper class. There followed more liberal theories of classes (see, e.g., Levy et al. 1973), but in none of these theories can a proper class be a self-member, and so Russell's paradox does not arise.

However, the introduction of classes seems merely to push the problem back. Still there is no extension for 'x = x,' or for the predicates 'class' and 'proper class.' And no class can serve as the domain of quantification over classes.

A more promising strategy, it would seem, is to develop a theory of extensions from scratch. In my view, the notions of *extension* and *set* (or *class*) are independent and mutually irreducible. We cannot reduce sets or classes to extensions, for extensions are essentially tied to predication, and sets and classes are not. (Given some natural assumptions, there are strictly more sets in the ZF hierarchy – and more classes – than there are predicates in, say, English.) And we cannot reduce extensions to sets or classes. No set can serve as the extension of 'set,' and no class can serve as the extension of 'class'; and there are, as we have seen, self-membered extensions, but no self-membered classes or ZF sets.

If we do develop a theory of extensions directly, we must of course find a way out of Russell's paradox for extensions. We saw in Section 1 that this paradox is best viewed as a semantical paradox, and that it shares structural similarities with the heterological paradox. All the better, then, if we can find a unified solution to this version of Russell's paradox, the paradoxes of definability, and the Liar paradoxes.

4 Three Paradoxes

In search of such a unified account, consider three paradoxes. First, suppose that I am confused about my whereabouts (I think I am in room 102), and I write on the board in room 101 the following denoting expressions:

(A) the ratio of the circumference of a circle to its diameter.
(B) the successor of 5.
(C) the sum of the numbers denoted by expressions on the board in room 101.

It is clear what the denotation of (A) and (B) are. But what is the denotation of (C)? Suppose (C) denotes k. Then the sum of the numbers denoted by expressions on the board is $\pi + 6 + k$. So (C) denotes $\pi + 6 + k$. So $k = \pi + 6 + k$. We are landed in a contradiction.

So we should conclude:

(6) (C) is pathological, and does not denote a number.

Now we can reason that (A) and (B) are the only expressions on the board that denote numbers. So we may conclude that *the sum of the numbers denoted by expressions on the board in room 101* is $\pi + 6$. Observe that in the previous sentence there occurs a token of the same type as (C), call it (C*). Unlike (C), (C*) is not pathological, and it does have a denotation. We may conclude:

(7) (C*) denotes $\pi + 6$.

How can two expressions – composed of exactly the same words with the same linguistic meaning – differ so dramatically in their semantic status?

Suppose next that I write on the board in room 101 these two predicates:

(E) moon of the Earth
(F) unit extension of a predicate on the board in room 101.

The extension of predicate (E) is a unit extension, and so it is a member of the extension of (F). What about the extension of (F)? Suppose first that it is a self-member. Then the extension of (F) has two members, so it is not a unit extension – and so it is not a self-member. Suppose second that it is not a self-member. Then the extension of (F) has just one member, so it is a unit extension – and so it is a self-member. Either way we obtain a contradiction.

So we should conclude that (F) is a pathological predicate that fails to have an extension. But if (F) does not have an extension, then in particular it does not have a unit extension. So the only *unit extension of a predicate on the board in room 101* is the extension of (E). We've just produced a token of the same type as (F), call it (F*). But unlike (F), (F*) has a well-determined extension (whose only member is the extension of (E)). Again we can ask: how is that these two expressions – composed of the very same words – differ in their semantic status?

Finally, consider the case of truth. If I write on the board in room 101 the following sentence:

(L) The sentence written on the board in room 101 is not true,

then I have produced a Liar sentence. We are landed in a contradiction whether we assume (L) is true, or not true. So we can conclude that (L) is semantically pathological. As we have seen, semantic pathologicality may be cashed out in a variety of ways – for example, perhaps (L) is gappy or unstable. But if (L) is pathological, then it is not true. That is, we may conclude:

(L*) The sentence written on the board in room 101 is not true.

And while (L) is pathological, (L*) is true. Again, the two sentences differ in semantic status, yet they are tokens of the same type.

5 A Contextual Approach

How should we resolve these paradoxes? In each case, we have the same phenomenon: a change in semantic value (in denotation, extension, or truth-value) without a change in linguistic meaning. Such a change suggests some pragmatic difference.

Consider the case of denotation, though what we say about this case carries over to the others. There are a number of differences between the context of (C) and the context of (C*). Beyond the familiar contextual parameters of speaker, time, and place, there are differences in *discourse position*, *intentions*, and *relevant information* as well. (C*) is produced at a later stage of the discourse, *after* it has been established that (C) is pathological. At this later stage, we reason in the light of (C)'s pathology – we may say that the context in which we produce (C*) is *reflective* with respect to (C). Intentions shift too. At the later stage, our intention is to treat (C) as pathological and see where this leads us. But I have no such intention at the first stage – my intention in producing (C) is to refer to expressions on the board next door. There is a corresponding shift in information: the information that (C) is pathological is available throughout the later stage of the reasoning, but it is not available to me when I first produce (C). These contextual differences all contribute to a crucial contrast between the contexts of (C) and (C*): the former is *unreflective with respect to (C)*, and the latter is *reflective with respect to (C)*.

If we accept the appropriateness of a pragmatic explanation, then we should expect to find a term occurring in (C) and (C*), and in (1) and (2), that is context-sensitive. When we inspect the terms occurring in these expressions, there seems to be only one candidate: the predicate 'denotes.' Accordingly, let us represent (C) by

(C) the sum of the numbers denoted$_C$ by expressions on the board in room 101,

where the subscript indicates that the use of 'denotes' in (C) is tied to (C)'s unreflective context of utterance.

To determine the denotation of (C), then, we must determine the denotation$_C$ of expressions on the board – that is, the denotations$_C$ of (A), (B), and (C). The conditions under which an expression denotes$_C$ is given by a denotation schema (analogous to the truth-schema):

s denotes$_C$ n iff p = n,

where instances of the schema are obtained by substituting for 'p' any referring expression, for 's' any name of this expression, and for 'n' any name of an individual. When we apply this C-schema to (C), we obtain a contradiction, and this leads to the conclusion (8), represented by:

(8) (C) does not denote$_C$ a number.

We go on to reason that (A) and (B) are the only expressions on the board that denote$_C$ numbers, since (C) does not. So we infer that the sum of the numbers denoted$_C$

by expressions on the board in room 101 is $\pi + 6$. In producing (C*) here, we have in effect repeated (C). But we have repeated (C) in a new context, a context that is reflective with respect to (C). We no longer provide denotation conditions via the C-schema. In this new reflective context – call it R – denotations are determined *in the light of (C)'s pathology*. That is, denotations are determined by the R-schema:

s denotes$_R$ n iff p = n.

And (C*) *does* have a denotation$_R$. Consider the biconditional:

(C*) denotes$_R$ k iff the sum of the numbers denoted$_C$ by expressions on the board in room 101 at noon 7/1/99 is k.

The right-hand side is true for $k = \pi + 6$, since we have established that (C) does not denote$_C$. And so we infer

(9) (C*) denotes$_R$ $\pi + 6$.

(C) and (C*) are semantically indistinguishable – the difference between them is a purely pragmatic one. It is a matter of the denotation schemas by which (C) and (C*) are given denotation conditions. At the first stage of the reasoning, (C) is assessed via the unreflective C-schema; at the second stage, (C*) is assessed via the reflective R-schema. Notice that if we assess (C) via the R-schema, we find that (C), like (C*), denotes$_R$ $\pi + 6$; and if we assess (C*) via the C-schema, we find that (C*), like (C), does not denote$_C$ a number. So the tokens of 'denotes' in (8) and (9) have different extensions: (C) and (C*) are not in the extension of 'denotes$_C$,' but both are in the extension of 'denotes$_R$.' So 'denotes' is a context-sensitive term that may shift its extension with a change in context – as it does in the move from (8) to (9).

We can give exactly parallel analyses of the cases of extension and truth. We take 'extension' and 'true' to be context-sensitive terms, and explain the difference between (F) and (F*), and between (L) and (L*), in terms of a change in evaluating schemas.

6 A Singularity Proposal

The question naturally arises: what is the relation between the unreflective and reflective stages? A possible response here is a Tarskian one: when we move from the first stage of the reasoning to the second, we push up a level of language. (For contextual accounts of truth that appeal to a hierarchy, see Parsons 1974; Burge 1979; Barwise and Etchemendy 1987; Gaifman 1988, 1992.) So, for example, the terms 'denotes$_C$' and 'denotes$_R$' belong to distinct levels, and the extension of 'denotes$_R$' properly contains the extension of 'denotes$_C$.'

We have already seen the difficulties that hierarchical accounts face. But a unified hierarchical account of reference, extension, and truth faces a special difficulty. It is the case of extensions that presents the problem. Extensions can be self-membered; for example, as we saw in Section 1, the extension of the predicate 'infinite extension' belongs to itself. According to the hierarchical approach, this predicate is of the form 'infinite extension$_\sigma$.' And the predicate itself is a predicate of $L_{\sigma+1}$ and not of L_σ. So the extension of this predicate is not a self-member – it contains extensions of predicates of L_σ only. The hierarchical account cannot accommodate self-membered extensions. A distinctive feature of extensions is regimented away.

But perhaps we can retain the contextual idea and jettison the hierarchy. This is the idea behind the *singularity theory* (see Simmons 1993, 1994). Occurrences of 'denotes,' 'extension,' and 'true' are to be *minimally* restricted, in accordance with the pragmatic principle of Minimality. Suppose, for example, you say " 'The square of 1' denotes 1." Here, your use of 'denotes' is quite unproblematic. Should (C) be excluded from its extension? According to Minimality, the answer is no – because there is no need to exclude it. We have seen that (C) denotes$_R$ $\pi + 6$ because the sum of the numbers denoted$_C$ by expressions on the board *is* $\pi + 6$. And for the same reason, (C) denotes$_N$ $\pi + 6$, where N is the context of your utterance, a context neutral with respect to (C).

If we adopt Minimality we respect a basic intuition about predicates. In general, if an individual has the property picked out by the predicate φ, then we expect that individual to be in the extension of φ. The more restrictions we place on occurrences of 'denotes' (or 'extension' or 'true'), the more we are at odds with this intuition. Minimality keeps surprise to a minimum.

So the present proposal identifies *singularities* of the concepts *denotes*, *extension*, and *truth*. For example, (C) is a singularity of 'denotes$_C$,' because it cannot be given denotation$_C$ conditions. Notice that (C) is a singularity only in a context-relative way – it is not a singularity of 'denotes$_R$' or 'denotes$_N$.'

No occurrence of 'denotes' or 'extension' or 'true' is without singularities. For example, consider again (7):

(7) (C*) denotes $\pi + 6$.

Consider the following perverse addition to (7):

(10) And so the number denoted by (C*), plus the number denoted by 'the square of 1,' plus the sum of the numbers denoted by phrases in this sentence, is irrational.

Given the context, the occurrences of 'denotes' in our continuation will be represented by 'denotes$_R$.' Consider the final definite description token in our utterance (beginning 'the sum of') – call this token (D). (D) is a singularity of 'denotes$_R$' – the R-schema cannot provide it with denotation conditions.

The example of (D) brings out the anti-hierarchical nature of the singularity proposal. Observe that we can reflect on (D), just as we earlier reflected on (C). In a suitably reflective context, we can conclude that (D) denotes $(\pi + 6) + 1$ – since the only denoting phrases in (10) that denote$_R$ numbers are the first two phrases, and these phrases denote $\pi + 6$ and 1 respectively. And by Minimality, (D) will have this denotation when assessed by any schema other than the R-schema. In particular, the token *does* denote$_C$ – it is *not* a singularity of 'denotes$_C$.' The C-schema does determine a denotation for it. On a Tarskian account, the extension of 'denotes$_C$' will be a proper subset of the extension of 'denotes$_R$.' According to the singularity proposal, neither extension includes the other.

Gödel once made the following tantalizing remark about the paradoxes:

> It might even turn out that it is possible to assume every concept to be significant everywhere except for certain 'singular points' or 'limiting points', so that the paradoxes would appear as something analogous to dividing by zero. Such a system would be most satisfying in the following respect: our logical intuitions would then remain correct up to certain minor corrections, i.e. they could then be considered to give an essentially correct, only somewhat 'blurred', picture of the real state of affairs. (Gödel 1944: 229)

I take the singularity proposal to be in the spirit of Gödel's suggestion. According to the present account, our intuitions about 'denotes' – and 'extension' and 'true' – are almost correct. It is only in pathological or paradoxical contexts that we may mistakenly suppose that certain phrases denote when they do not – and in such cases our applications of 'denotes' require only *minimal* corrections. We retain a single denotation predicate which undergoes minimal changes in its extension according to context. There is no wholesale revision of the notion of denotation; no division of 'denotes' into infinitely many distinct predicates; no splitting of everyday English into an infinite hierarchy of languages.

7 Universality

It is of course beyond the scope of this chapter to provide a formal theory of singularities (see Simmons (1993) for a singularity theory of truth). But suppose we had such a formal theory $\mathcal{L}_T$ for an object language $\mathcal{L}$ containing the context-sensitive predicate 'denotes,' or 'extension,' or 'true.' Won't we now face the familiar objection, that $\mathcal{L}_T$ is a metalanguage for $\mathcal{L}$, and the hierarchy is inevitable? Moreover, since we may regard $\mathcal{L}_T$ as a classical formal language, it will be subject to Tarski's theorem. So the semantical predicates for $\mathcal{L}_T$ will be contained in a further metalanguage, which in turn cannot contain *its* semantic predicates. From $\mathcal{L}_T$, then, a whole hierarchy of languages is generated.

But the singularity account is not without resources here. The context-sensitive predicate 'denotes' applies to any denoting phrase of $\mathcal{L}_T$ as long as that phrase is not identified as a singularity – and similarly for any denoting phrase at any level of the ensuing hierarchy. No language of the hierarchy is a metalanguage for $\mathcal{L}$ – 'denotes' applies to phrases of all levels. (Parallel remarks can be made about 'extension' and 'true.') The scope of 'denotes' is as close to universal as it can be.

According to Tarski, natural languages are "all-comprehensive" and "universal":

> The common language is universal and is intended to be so. It is supposed to provide adequate facilities for expressing everything that can be expressed at all, in any language whatsoever; it is continually expanding to satisfy this requirement. (Tarski 1969: 89)

It is the apparent universal character of natural language that both generates semantical paradoxes and makes them so difficult to solve. Any semantic account of 'denotes,' 'extension,' or 'true' will just be more English, and so the stage is set for a revenge Liar. Whether the singularity account, or some other, can do sufficient justice to this feature of natural language cannot be settled here. But the challenge remains. At root, semantical paradox and the problem of universality are one and the same.

References

Barwise, Jon and Etchemendy, John (1987) *The Liar*. Oxford: Oxford University Press.

Boolos, George (1971) The iterative conception of set. *Journal of Philosophy*, 68, 215–32; reprinted in Putnam and Benacerraf (1983) 486–502.

Burge, Tyler (1979) Semantical paradox. *Journal of Philosophy*, 76, 169–98; reprinted with a postscript in R. L. Martin (ed.) (1984) 83–117.

Feferman, Solomon (1982) Towards useful type-free theories, I. *Journal of Symbolic Logic*, 49, 75–111; reprinted in R. L. Martin (ed.) (1984) 237–87.

Gaifman, Haim (1988) Operational pointer semantics: solution to self-referential puzzles I. *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, (eds.) M. Vardi and Morgan Kaufman. Los Altos, California, 43–60.

Gaifman, Haim (1992) Pointers to truth. *Journal of Philosophy*, 223–61.

Gödel, Kurt (1944) Russell's mathematical logic. In P. A. Schilpp (1944) 123–53.

Gupta, Anil (1982) Truth and paradox. *Journal of Philosophical Logic*, 11, 1–60; reprinted in R. L. Martin (ed.) (1984) 175–235.

Gupta, Anil and Belnap, Nuel (1993) *The Revision Theory of Truth*. Cambridge, MA: MIT Press.

Herzberger, Hans (1982) Notes on naive semantics. *Journal of Philosophical Logic*, 11, 61–102; reprinted in R. L. Martin (ed.) (1984) 133–74.

König, Julius (1905) On the foundations of set theory and the continuum problem. In van Heijenoort (1967) 145–9.

Kripke, Saul (1975) Outline of a theory of truth. *Journal of Philosophy*, 72, 690–716; reprinted in R. L. Martin (ed.) (1984) 53–81.

Levy, A., Fraenkel, A. A. and Bar-Hillel, Y. (1973) The role of classes in set theory. In Muller (1976) 173–215; first published as Chapter II, Sec. 7 of *Foundations of Set Theory*. Amsterdam: North-Holland.

Maddy, Penelope (1983) Proper classes. *Journal of Symbolic Logic*, 48, 113–39.

Martin, Robert L. (ed.) (1984) *Recent Essays on Truth and the Liar Paradox*. Oxford: Oxford University Press.

McGee, Vann (1990) *Truth, Vagueness, and Paradox*. New York: Hackett.

Parsons, Charles (1974) The liar paradox. *Journal of Philosophical Logic*, 3, 381–412; reprinted with a postscript in R. L. Martin (1984) 9–45.

Priest, Graham (1979) The logic of paradox. *Journal of Philosophical Logic*, 8, 219–41.

Priest, Graham (1984) Logic of paradox revisited. *Journal of Philosophical Logic*, 13, 153–79.

Quine, W. V. (1937) New foundations for mathematical logic. In Quine (1953) 80–101.

Quine, W. V. (1953) *From a Logical Point of View*. New York: Harper Torchbooks.

Ramsey, Frank (1925) *The Foundations of Mathematics*. London: Routledge & Kegan Paul.

Richard, Jules (1905) Les principes des mathématiques et le problème des ensembles. In *Revue générale des sciences pures et appliquées*, 16, 541. Also in *Acta Mathematica*, 30 (1906), 295–6. (English translation in van Heijenoort (1967) 143–4.)

Russell, Bertrand (1903) *The Principles of Mathematics*. Cambridge: Cambridge University Press.

Russell, Bertrand (1908) Mathematical logic as based on the theory of types. *American Journal of Mathematics*, 30, 222–62; reprinted in van Heijenoort (1967) 150–82.

Schilpp, P. A. (ed.) (1944) *The Philosophy of Bertrand Russell*. La Salle, IL: Open Court.
Simmons, Keith (1993) *Universality and the Liar: An Essay on Truth and the Diagonal Argument*. Cambridge: Cambridge University Press.
Simmons, Keith (1994) Paradoxes of denotation. *Philosophical Studies*, 76, 71–104.
Tarski, Alfred (1944) The semantic conception of truth. *Philosophy and Phenomenological Research*, iv, 341–76; reprinted in Leonard Linsky (ed.), *Semantics and the Philosophy of Language*. Chicago: University of Illinois, 1952, 13–47.
Tarski, Alfred (1969) Truth and proof. *Scientific American*, 220, 63–77.
Van Heijenoort, Jean (ed.) (1967) *From Frege to Gödel: A Source Book in Mathematical Logic*. Cambridge, MA: Harvard University Press.
Von Neumann, John (1925) An axiomatization of set theory. In Van Heijenoort (1967) 394–413.