Chapter 4 Paradoxes of Truth, Satisfaction and Membership

This chapter investigates the relationships between paradoxes of truth, satisfaction and membership. By classifying the paradoxes of truth, satisfaction and membership, an infinite variety of paradoxes are brought under a potentially finite classification. We are then in a better position to assess whether these paradoxes are of a uniform type. This is important because it is a key tenet of those who advocate a uniform solution. Historically, there have been two orthodoxies about types of these paradoxes, one deriving from Russell [1908] that holds they are of a uniform type, and another from Peano [1906] and Ramsey [1925] that holds the semantic paradoxes are of a different type from the set-theoretical paradoxes. I argue that there are indeed two types, but that the paradigm cases of these are two different paradoxes of satisfaction, including a new one I will introduce.

4.1 Introduction

In this chapter, I map a classification of the family of Liar paradoxes and *hypodoxes* identified in the previous chapter into the Satisfaction family. The Liar family of paradoxes includes the Eubulidean Liar, quantified and unquantified variations of the Epimenides and Curry paradoxes, the Eldridge-Smith paradox (the ESP), quantified and unquantified circular Liar-like paradoxes, Sorensen's queue and Yablo's paradoxes.[1] Moreover, I made and investigated the conjecture in the previous chapter that each of the paradoxes of the Liar family has a *hypodoxical* dual. I will show that each of these paradoxes and hypodoxes can be mapped to a paradox or hypodox that

[1] With regard to naming conventions, there are few. I endeavour to respect common usage, I suggest 'the Epimenides paradox', 'Curry's paradox' or 'Curry paradoxes', 'Russell's paradox' and 'Grelling's paradox' are common usage. I prefer to use 'the ... paradox' when referring to a category within a family of paradoxes. Each category of paradoxes has instances, to one of which I may refer using 'a ... paradox'. For example, Tarski used the letter '*c*' to name his example of the Liar; whereas I will talk about *my favourite sentence*, which is also a Liar sentence. By contrast, 'Russell's paradox' might be taken to name a particular paradox, not a type. 'Grelling's paradox' has 'the Heterological paradox' as a synonym. But I am not defending my nomenclature. I am just *explaining* it. Later, I will defend a distinction.

uses the satisfaction relation (or its converse, the 'is true of' relation). However, there are still more satisfaction paradoxes.

There are *prima facie* two branches within the family of satisfaction paradoxes. Members of one branch are more closely analogous to the Liar family, and members of another branch are more closely analogous to Russell's paradox. Thus, it initially seems that there are twice as many paradoxes of satisfaction as there are Liar paradoxes, and, more to the point, it seems they are of two types.

Section two provides some background to this debate. I illustrate the two types of paradox of satisfaction in section three. I discuss naive theories, and characterise the distinction between different types of paradoxes in section 4. In section 5, I discuss the Property paradox, Heterologicality and Quine's Liar in relation to the two types of paradox. In section 6, I map Liar-like paradoxes to satisfaction paradoxes of the first type, show (in most cases that) there is a second type of satisfaction paradox and map that to a set-theoretic paradox. I will present some derivations, but not dwell on them. I conclude by discussing the significance of the two types of paradox.

4.2 Some Background on Orthodox Theories about Types of these Paradoxes

In *The Structure of the Paradoxes of Self-reference*, Priest [1994] argued that a high-level uniform description of the paradoxicality of semantic, set-theoretic and definitional paradoxes implied a uniform solution. (This solution might not be paraconsistent. Goldstein [1999], for example, maintains that Yablo's paradox is circular after all and that his development of the *no-statement* approach to the Liar is uniform with the non-existent set approach to Russell's paradox.) Priest is arguing against an orthodox view that the semantic and set-theoretic paradoxes are of different types. Priest stresses a formal analogy and abstracts away from the use of different notions in the paradoxes:

> ... the correct level of abstraction for an analysis of the paradoxes of self-reference is not one which depends upon the presence of certain words ("set", "true", etc.), but the level of the underlying structure that generates and causes the contradictions...
>
> [Priest 2000, p. 125]

> ... if one wants to draw a fundamental distinction, this ought to be done in terms of the *structure* of the different paradoxes. Ramsey's distinction depends on the relatively superficial fact of what vocabulary is used in the paradoxes, and, in particular, whether this belongs to mathematics properly so called. But worse, this is a notoriously shifting boundary....
>
> [Priest 1994, p. 26]

For over a century, intuitions have polarised over whether or not the semantic and set-theoretic paradoxes are of a uniform type. Russell's [1908, p. 225] intuitions were that there was a common pathology, *the Vicious Circle Principle*. Ramsey [1925 / 1931, p. 20-21], taking up a distinction made originally by Peano, believed that the paradoxes were of two distinct types. Although Ramsey gave a sophisticated articulation of the distinction, it was partly based on set-theoretic notions being amenable to formal definition, and semantic notions not being amenable. This basis is no longer valid. Nevertheless, intuitions are still polarised. Accordingly, Ramsey has been re-interpreted, as discussed below.

Mendelson [1979, p. 3] represents the distinction as based on the set-theoretic paradoxes relying on *fewer*, or at least more fundamental, logical and mathematical notions.

> The logical [*sic*] paradoxes involve only notions from the theory of sets, whereas the semantic paradoxes also make use of concepts like "denote", "true", "adjective", which need not occur within our standard mathematical language. For this reason, the logical paradoxes are a much greater threat to a mathematician's peace of mind than the semantic paradoxes.
>
> [Mendelson 1979, p. 3]

The validity of Peano's and Ramsey's distinction insulates foundational mathematics from paradox, if the set-theoretic paradoxes are resolved.

Quine bases Peano's and Ramsey's distinction on the different concepts involved:

> The notions of denotation, truth, and specifiability must be subjected to some sort of intuitively unanticipated restriction, in the light of these [semantic] paradoxes, just as class existence must in the light of

> Russell's paradox and others. But the semantic paradoxes are of no concern to the theory of classes.
>
> [Quine 1969, p. 255]

Quine acknowledges a formal analogy between some paradoxes but does not rely on it, and instead stresses the use of distinct notions:

> Russell's antinomy bears a conspicuous analogy to Grelling's antinomy of 'not true of self', which in fact it antedates. But Russell's antinomy does not belong to the same family as the Epimenides antinomy and those of Berry and Grelling. By this I mean that Russell's antinomy cannot be blamed on any of the truth locutions, nor is it resolved by subjecting those locutions to subscripts. The crucial words in Russell's antinomy are 'class' and 'member', and neither of these is definable in terms of 'true', 'true of', or the like.
>
> [Quine 1962/1976, p. 11]

Related to Quine's view is the view that the paradoxes are to be addressed by replacing the naive notions with other intuitions about truth, satisfaction and membership. The paradoxes can be used as *reductios* of our naive notions of truth, satisfaction and membership. Since each of these naive notions is inconsistent, they are to be replaced by other (consistent) notions of truth, satisfaction and membership, such as the iterative conception of set. These other (consistent) notions need not bear analogies to each other.

The Peano-Ramsey distinction makes it reasonable to provide independent solutions to the paradoxes of set-theory and those of semantics. The paradoxes of set-theory can be resolved by denying the existence of the Russell set and by using a different conception of set – the iterative conception. Intuitively analogous approaches to the Liar, for example, are to deny the Liar makes a statement or expresses any proposition and to use a *grounded* notion of truth. The analogies may be interesting; but they are of no significance if the Peano-Ramsey distinction is correct. If the distinction were ill-founded and the paradoxes are of a fundamentally uniform nature, then analogies are not only significant, but one might expect a uniform solution.

I will distinguish the satisfaction paradox that Russell's paradox resembles from another satisfaction paradox more closely resembling the Eubulidean Liar. *Prima*

facie, the different types of satisfaction paradox provide a basis for a distinction, although the line has shifted from where Peano and Ramsey drew it.

4.3 Two Branches of the Satisfaction Family

In this section, I show that there appear to be more paradoxes of satisfaction than Liar-like paradoxes as the satisfaction paradoxes fall into two groups, one of which corresponds to the Liar-like paradoxes. The members of the other branch of the satisfaction paradoxes can be mapped into the set-theoretic paradoxes.

Here is the Liar:

> My favourite sentence just happens to be 'My favourite sentence is not true'. So, my favourite sentence is true if and only if it is not true.

And here is a satisfaction paradox that parallels it:

> My favourite predicate just happens to be 'does not satisfy my favourite predicate'. So, Knossos satisfies my favourite predicate if and only if Knossos does not satisfy my favourite predicate.

I draw this parallel formally below, where 'Sat' is a two-or-more-place satisfaction-predicate. The Liar uses the T-schema, a canonical version in this derivation, and the unsatisfied paradox uses a similar schema for satisfaction.

	The (Eubulidean) Liar:	
1	$a = \langle \sim Ta \rangle$	Premise
2	$T\langle \sim Ta \rangle$ iff $\sim Ta$	1, T-schema
3	Ta iff $\sim Ta$	2, =E (Leibniz's law)

	The Unsatisfied Predicate Paradox[2]:	
1	$p = \langle \sim(x \text{ Sat } p) \rangle$	Premise
2	t Sat $\langle \sim(x \text{ Sat } p) \rangle$ iff $\sim(t \text{ Sat } p)$	1, Satisfaction schema
3	t Sat p iff $\sim(t \text{ Sat } p)$	2, =E

[2] I devised this satisfaction paradox by analogy with the Eubulidean Liar in 2003. I believe it is original. I presented it in a paper corresponding to this chapter at the 2005 Australasian Association of Philosophy conference.

Whether each identity premise in line 1 of each argument is empirically true or is true by stipulation makes no difference to the argument's seeming valid. So long as one grants that line 1 could be true the paradox is adequately represented for present purposes. As discussed in Chapter 3, similar arguments apparently differ at line 1, for example by using an indexical (e.g. 'This sentence is false' and 'does not satisfy this predicate') or a biconditional that is obtained using arithmetization.

Just as the Liar uses an instance of the (canonical) T-schema, the Unsatisfied Paradox uses an instance of the (canonical) Satisfaction Schema:

> t Sat $\langle \mathcal{P}x \rangle$ iff $\mathcal{P}(t/x)$
>
> where '$\mathcal{P}x$' represents an open sentence with a free variable, loosely speaking, here and throughout, a predicate.[3]

The statements and predicates embedded in the above identity premises are paradoxical, not the identities themselves. We can imagine circumstances that would make the above identities true. At least, that is, if we take the identities to be co-referring terms for a sentence in the case of the Liar, and a predicate in the case of the Unsatisfied paradox. The Liar sentence, 'My favourite sentence is not true' is grammatically well formed, and I favour it. The Unsatisfied predicate 'does not satisfy my favourite predicate' is a grammatically well-formed predicate, and, I favour it.

Most sentences do not pose Liar-like issues for using the truth-predicate, yet whether or not any thing satisfies the Unsatisfied Predicate is problematic. So, the Unsatisfied Predicate paradox is not solved by keeping the predicate itself out of its own range of applicability, although keeping the Liar sentence out of the range of its truth predicate is the essential strategy of some orthodox solutions to the Liar.

There is an analogous argument using the set abstraction schema:

$$z \in \{x{:}Px\} \text{ iff } Pz$$

This argument is a *reductio*, proving the set identity in line 1 logically false. Unlike the identity for a, one cannot imagine circumstances that would make the identity for q below true. No set is identical with its own complement:[4]

---

[3] In this chapter, I restrict discussion to one-place open sentences. In general, $(t_1, t_2, \ldots)$ sat $\langle \mathcal{P}(x_1, x_2, \ldots)\rangle \equiv \mathcal{P}(t_1/x_1, t_2/x_2, \ldots)$

[4] I naively omit that it could be true if nothing exists, in which case a *free* logic could avoid the *reductio*.

	A non-set:	
1	$q = \{x: x \notin q\}$	Premise
2	$t \in \{x: x \notin q\}$ iff $t \notin q$	1, Abstraction schema
3	$t \in q$ iff $t \notin q$	2, Identity

There is *prima facie* another type of satisfaction paradox. In contrast to the Unsatisfied Predicate paradox, consider the following argument towards a version of Grelling's paradox.

	Grelling's paradox:	
1	$r = \langle\sim(x \text{ Sat } x)\rangle$	Premise
2	$r \text{ Sat } \langle\sim x \text{ Sat } x\rangle$ iff $\sim(r \text{ Sat } r)$	1, Satisfaction schema
3	$r \text{ Sat } r$ iff $\sim(r \text{ Sat } r)$	2, Identity

We can render this loosely in English as follows: Suppose Grelling's favourite predicate is 'does not satisfy itself', which is true of an expression that does not satisfy its own condition; but then Grelling's favourite predicate satisfies itself iff it does not satisfy its own condition. This is analogous to Russell's paradox (below), but does not appear to have a counterpart Liar, as the Unsatisfied Predicate paradox is the counterpart of the Liar using the Satisfaction Schema.

	Russell's paradox:	
1	$w = \{x: x \notin x\}$	Premise
2	$w \in \{x: x \notin x\}$ iff $w \notin w$	1, Abstraction schema
3	$w \in w$ iff $w \notin w$	2, Identity

The *Unsatisfied Predicate* is distinct from *Grelling's*. Here are two differences: 1) In the former, we can use any term to derive a contradiction. 2) In the latter, the identity premise is a mere abbreviation. We can use an angle bracket name instead of '*r*' in an instance of the satisfaction schema for this predicate, and derive a contradiction in one step as below – a one step argument from no premises:

$$\langle\sim x \text{ Sat } x\rangle \text{ Sat } \langle\sim x \text{ Sat } x\rangle \equiv \sim(\langle\sim x \text{ Sat } x\rangle \text{ Sat } \langle\sim x \text{ Sat } x\rangle)$$

We can illustrate the first difference between the *Unsatisfied Paradox* and *Grelling's paradox* with indexicals as well. Compare the predicate 'does not satisfy

this predicate' against 'does not satisfy itself'. To satisfy the former predicate is not to satisfy it. Nothing can consistently satisfy it or not satisfy it. Whereas for something to satisfy the latter predicate is just for it not to satisfy itself. 'Long', for example, would satisfy the latter predicate as 'long' is not long. The latter predicate has a specific problem as to whether or not it satisfies itself.

While arithmetization would reduce the identity premise to a theorem, the Liar-like paradoxes still require such an identity; but, as I have noted, the identity is a mere abbreviational convenience for the Unsatisfied Grelling's and Russell's paradox. So here is a difference, the Liar requires the identity premise, but Grelling's is derivable without an identity (premise or theorem) as an instance of the Sat-schema.

The table below summarises the classification of these paradoxes.

Type-1	Type-2
(Eubulidean) Liar (4th Century BCE)	
$a = \langle \sim Ta \rangle$ $T\langle \sim Ta \rangle$ iff $\sim Ta$	
The Unsatisfied paradox	Grelling's
$p = \langle \sim(x \text{ Sat } p) \rangle$ t Sat $\langle \sim(x \text{ Sat } p) \rangle$ iff $\sim(t \text{ Sat } p)$ t Sat p iff $\sim(t \text{ Sat } p)$	$r = \langle \sim(x \text{ Sat } x) \rangle$ r Sat $\langle \sim(x \text{ Sat } x) \rangle$ iff $\sim(r \text{ Sat } r)$ r Sat r iff $\sim(r \text{ Sat } r)$
A non-set:	Russell's set
$q = \{x: x \notin q\}$ $z \in \{x: x \notin q\}$ iff $z \notin q$ $z \in q$ iff $z \notin q$	$w = \{x: x \notin x\}$ $w \in \{x: x \notin x\}$ iff $w \notin w$ $w \in w$ iff $w \notin w$

I note that there are also dual hypodoxes for each of these paradoxes. They are each obtained by the external negation of the paradoxical expressions. In each case their semantic value, their truth value or extension, is under-determined in some respect, as indicated. There is a lack of any basis to decide whether the Truth-teller is true or not.

Similarly, there is a lack of anything to determine whether anything satisfies the hypodox corresponding to my new paradox of satisfaction. The hypodoxical issue for autological is localised to autological itself, as nothing determines whether autological satisfies itself.

Although the identity in the table below for the putative set q_e looks like a definition, it is circular. As there is no independent reason to believe the set q_e exists, and, although formally analogous to e and p_e it does not succeed as any sort of stipulation. There may, nevertheless, be independent reasons for the truth of identities of the form of e and p_e.

Type-1	Type-2
The Truth-teller	
$e = \langle \mathrm{T}e \rangle$ The sentence e is naively true or not true, but nothing determines which. Contingent facts or conventions may make an identity of this form true.	
The Satisfied hypodox	Grelling's Autological Hypodox
$p_e = \langle (\mathrm{x\ Sat}\ p_e) \rangle$ Naively, anything in the domain either satisfies p_e or not, but nothing determines which in any case. Nevertheless, contingent facts or conventions may make an identity of this form true.	$r_e = \langle (\mathrm{x\ Sat\ x}) \rangle$ Does r_e satisfy r_e or not?
Another non-set, (and non-hypodox):	Russell's hypodoxical set
$q_e = \{\mathrm{x}: \mathrm{x} \in q_e\}$ This cannot be a contingent identity. As a definition, it is circular, and fails to define a particular set, any set would satisfy this definition. It cannot succeed as any other kind of stipulation, for the reference of q_e is indeterminate.	$w_e = \{\mathrm{x}: \mathrm{x} \in \mathrm{x}\}$ Is w_e a member of w_e or not?

These identities look like definitions and all the Type-1s in the left-hand side of the table may appear to be circular because the term being defined occurs in both sides of the identity statements. In fact the identity of e and p_e need not be interpreted as stipulations, there may be other reasons to accept them as sound premises.

The identity for the Satisfied Hypodox is merely specifying co-reference of the name of the predicate, p_e, with the grammatical predicate of the form within the angle

brackets. The identity would fail as an extensional definition if it were picking out the extension of the predicate; however, it is not, that is, not as standardly construed. If the angle brackets are taken to represent quotes for example, a standard construal of quotation names is as proper names; and the inclusion of the 'p_e' in the quotation name is just part of a meaningless string. Another interpretation of quotation names which attributes structure to the quotation names might consider that the identity, as a definition, is circular. However, this does not account for two other ways in which identities of this form may be true.

Nevertheless, the Truth-teller and the Satisfied hypodox may be contingently true, or even provable based on certain conventions or stipulations for canonical naming of expressions. The predicate p_e just happens to be my favourite hypodoxical predicate. Alternatively, arithmetization could be used to provide a canonical name for an identity of this form.

Although syntactically analogous, the identity statement for the set q_e is interpreted as picking out a set with reference to a predicate, such that the name is co-referential with the extension of the predicate. However, as the predicate that is used in specifying q_e, '$\in q_e$', uses the name of the set itself, as a definition, the identity is clearly circular. Any set would satisfy this condition. While the condition of q could not be satisfied by anything, the condition of q_e is always true, whatever is chosen as the extension of q_e. The putative identity for q_e cannot be contingently true, it cannot be a definition, and as some other form of stipulation it does not succeed.

In the Type-2 hypodoxes on the right-hand side of the table, there is no question of circularity because, taken as definitions, the *definiendum* does not occur in the *definiens*.

A distinction between the Liar and Grelling's has been considered before, and previously rejected. I give some account of previous considerations in the following.

Quine [1962, pp. 4-11] emphasises the similarities between Russell's and Grelling's, even though Quine does not support a uniform solution. Quine constructs what he puts forward as a Liar paradox more analogous to Grelling's. He uses a predicate 'yields a falsehood when appended to its own negation'. This does indeed do away with the need for a premise; for the following is analytically true:

> 'yields a falsehood when appended to its own negation' yields a falsehood when appended to its own negation

means

> " 'yields a falsehood when appended to its own negation' yields a falsehood when appended to its own negation" is false.

If Quine is correct and the paradox associated with the above equivalence does not need an identity premise, this might mean Quine's Liar might fill the vacant top right hand cell of the table of paradoxes. I return to discuss Quine's Liar in a subsequent section.

Despite his observations of analogy and disanalogy, Quine finds the source of the semantic paradoxes in truth locutions but that the set-theoretic paradoxes stem from existence assumptions.

Goldstein [2000, p. 66] rehearses a mapping from a definition of the Russell set to Grelling's predicate, 'is heterological'. He compares the definition of Russell's set with a definition for a non-set, like q above, but does not contrast them. He argues instead for a direct analogy between the heterological paradox, Russell's paradox and Liar.

Visser makes similar observations to Quine's:

> Grelling's Paradox shows clearly that Liar-like phenomena occur even in the absence of anything that can be described as self reference of a sentence, here self *application* takes the place of self reference. It might be argued that Grelling's paradox is different from the Liar ..., but nobody would want to argue thus against the following example [of a Circular Liar].
>
> [Visser 2004, p. 156]

I have presented two *prima facie* different types of satisfaction paradox, *the Unsatisfied paradox* and Grelling's. The former is more like the Liar. The latter is more like Russell's. Yet Quine, Visser and others have considered a distinction between Grelling's and the Liar paradox and hitherto rejected it. However, I am motivated by the intuitive contrast between Grelling's and the *Unsatisfied paradox.* Neither Quine, Goldstein nor Visser discusses or even considers the *Unsatisfied Paradox* as distinct from Grelling's paradox. I believe the *Unsatisfied Paradox* is original and that they were unaware of it. The *Unsatisfied* predicate is paradoxical no matter what we assume satisfies it, whereas Grelling's predicate is singularly paradoxical when applied to itself. The *Unsatisfied Paradox* requires an identity premise (or the like), whereas Grelling's paradox does not. Nevertheless, if Quine's Liar, which does not require an identity premise or identity theorem, is *just* an

economical derivation of a Liar paradox in this respect, then perhaps Grelling's paradox is *just* an economical derivation of a Satisfaction paradox. Before analysing Quine's paradox, I want to briefly discuss the schemata that characterise these three families of paradoxes and the two types among their paradoxes.

4.4 Naive Theories and their Discontents: Type-1 and Type-2

My present task is to classify the paradoxes, not to solve them. Accordingly, I work with so-called *naive theories* of truth, satisfaction and sets. Naively, we seem to have the idea that a one-place predicate *divides* things into two collections, those to which the predicate truly applies and those to which it does not, those in its extension and those not. A predicate's *Range of Applicability* (RA) is the maximal collection so divided. The truth-predicate is just another predicate making such a division. So are one-place satisfaction predicates, such as the predicate in the Unsatisfied paradox. The satisfaction relations and membership relations range over maximal collections of ordered pairs; but as they are used reflexively in Grelling's paradox and Russell's paradox, they are considered as ranging over individuals. (I will return to this point in the next chapter.)

I dare say I am working with *totally naive* theories, as I shall assume that there is a totality of things for each predicate's RA. (An alternate pre-theoretic *intuition* assumes that not every predicate has a totality of things to which it applies, but we will assume *total naivety*.) A theory would be *absolutely totally naive* if it assumed that the RA for every predicate is the same totality, but none of our paradoxes need that assumption.

> So, it would seem, for any predicate [φ] there is a set of all and only those things to which it applies (as well as a set of just those things to which it does not apply).... So (the universal closure of) $\ulcorner\exists y\,(\mathrm{S}y \wedge \forall x(x \in y \leftrightarrow \varphi))\urcorner$ should express a truth about sets (if no occurrence of "y" in φ is free).
>
> We call the theory whose axioms are the axiom of extensionality [that any two sets with exactly the same members are the same set] ... and all formulas $\ulcorner\exists y\,(\mathrm{S}y \wedge \forall x(x \in y \leftrightarrow \varphi))\urcorner$ (where "y" does not occur free in φ) *naive set theory*.
>
> [Boolos 1971 / 1998, p. 14...15].

Naive set theory is characterised by the abstraction schema. I note incidentally that Boolos' definition of *naive set theory* would summarily rule out the set q being stipulated as $q = \{x: x \in q\}$. Boolos uses 'y' as a variable ranging over sets but then uses it schematically to rule out any name of the set from being a component of the predicate, φ, which is used to specify that same set. I have been so naive as to leave off such a restriction. We label the extension of a predicate, $\mathcal{P}$, as '$\{x: \mathcal{P}x\}$', which is an *abstract name* of the set, and provide thereby a canonical means to form a name of the set from the predicate its members satisfy. Utilising this form of the name of a set, the principle of set abstraction is schematised as:

$z \in \{x:\mathcal{P}x\}$ iff $\mathcal{P}z$, where '$\mathcal{P}$' stands for any predicate.[5]

So, we have the following analogous schemata:

Abstraction schema: $z \in \{x:\mathcal{P}x\}$ iff $\mathcal{P}z$

Satisfaction schema: t sat $\langle\mathcal{P}x\rangle$ iff $\mathcal{P}(t/x)$

T-schema: $T\langle\mathcal{Q}\rangle$ iff $\mathcal{Q}$

Their implicit or explicit uses characterises our three families of paradoxes.

In the previous section, I introduced two types of paradoxes involving the satisfaction schema. One is paradoxical for every member of its RA; the other just if it itself is in its own RA. One needs an identity in its derivation; the other does not. The Liar appears to be a Type-1 paradox, while Russell's paradox appears to be a Type-2 paradox. The rest of this section is concerned with characterising these types of paradox.

4.4.1 Informal Semantic Criteria

I made some general observations about what distinguishes these types of paradox at the end of the last section; but one might expect the distinction to be guided by a semantic distinction. However, the *prima facie* semantic guidelines are too loose to characterise each type of paradox. Paradoxes of Type-1 involve self-reference,

---

[5] We can substitute '$\mathcal{P}x$' for 'φ' in Boolos' form of the abstraction schema to derive:

$\exists y\,(Sy \wedge \forall x(x \in y \leftrightarrow \mathcal{P}x))$

Then the axiom of extensionality guarantees uniqueness; so that we can label the set of things satisfying '$\mathcal{P}$', for which we can conventionally use $\{x:\mathcal{P}x\}$, and thus we have (with relettering):

$\forall z(z \in \{x:\mathcal{P}x\} \leftrightarrow \mathcal{P}z)$

or schematically:

$z \in \{x:\mathcal{P}x\} \leftrightarrow \mathcal{P}z$

circular-reference, or *infinite reference* as in referring to an infinite list of statements. Paradoxes of Type-2 involve reflexivity. The reflexive application of the membership condition for a set is problematic *only* in some cases. Russell's set is self-membered iff it is not. The reflexive application of satisfaction to a property, likewise, gives us the Property paradox, as explained in the next section. And this hardly exhausts the variety of Type-2 paradoxes. This concept of reflexivity must be broadened so that it is achievable in a circular manner or even in an infinite manner as we shall see. And when I talk about reflexivity this is what I mean.

Shortly, I will give proof-theoretic criteria for distinguishing the two types, which are sufficient for this chapter's purposes, but clarifying the semantic distinction is so desirable that I will attempt it here.

Generally and loosely speaking, problems of both types arise when the process of determining whether or not something satisfies a predicate cannot proceed in some normal way for lack of some independent criterion.

There are a number of intuitions about the types of paradoxes. Let me discuss five below.

Firstly, there is the intuition that for some paradoxes like the Liar, the existence of the problematic sentence or whatever is independent of the predicate concerned, whereas for other paradoxes like Russell's the existence of the set is dependent on the predicate concerned. Put another way, the Liar sentence is naively in the range of applicability of the truth predicate, whereas we do not know whether Russell's set is in the range of applicability of its predicate or not. In so far as this intuition is just about the Liar and Russell's paradoxes, it is unclear whether it motivates the semantic set-theoretic distinction any more than the distinction I have sketched.

This might seem to align with folk beliefs that philosophers' concerns over the Liar and semantic paradoxes are merely linguistic issues, whereas mathematicians' concerns with set theoretic paradoxes are more fundamental or at least not merely linguistic. Paradoxes like the Liar are merely linguistic because the truth of its premise arises from linguistic facts or stipulation, whereas paradoxes like Russell's are more concerned with reality, in particular ontology. There is a material issue with this view, because paradoxes like Curry's appear to arise from contingent non-linguistic facts in both naive set-theory and semantics. Nevertheless, on this view the import of the semantic paradoxes are taken to be purely philosophical, whereas the set theoretic paradoxes have import for foundational mathematics. This intuition at first appears to align well with a distinction between the semantic and set theoretic paradoxes.

Grelling's predicate is linguistic, that is true; so, by this intuition it ought to be classed with the other semantic paradoxes, and not with Russell's. Nevertheless, Grelling's paradox is formally analogous to the Property paradox, which I will discuss in the next section. But whether there is such a property is an ontic claim. The Property paradox belongs with Russell's going by this view; yet the Property paradox and Grelling's are formally the same paradox; so our distinction cannot be made purely formally. This is a consequence of such a criterion. Perhaps this consequence is inconvenient, but it is by no means a *reductio* of this linguistic *versus* non-linguistic criterion.

However, if this criterion is applied to the set q it seems to give the wrong determination. For the *reductio* of $q = \{x: x \notin q\}$ would seem to be an ontic paradox, not just a *reductio* of this stipulated identity. I think the argument from $q = \{x: x \notin q\}$ to a contradiction is a *reductio* of that stipulation, not a paradox. One would be more than just naive to suppose the existence of the set q depends on a predicate. An attempted definition fails because no set can satisfy this definition. If one makes a distinction as I have sketched, then q is formally like the Liar, not Russell's paradox, and the reason there is no paradox about q is that the argument to a contradiction is simply a *reductio* of the stipulated identity. Indeed, I think that what distinguishes q from the Liar paradox is not anything formal, but an independent reason to accept the truth of the identity premise in the case of the Liar.

Indeed the Liar *paradox* itself should be distinguished from *reductio* arguments with the same form. Visser is a source for such a formal Liar-look-a-like:[6]

1.	l iff $\sim l$	Premise, supposed true by stipulation
2.	|_____ l	Assumption
3.	| $\sim l$	1 defn., 2 MP
4.	$\sim l$	2-3 ~I
5.	l	1 defn., 4 MP
6.	l & $\sim l$	4, 5 &I

Some people such as Visser conditionally accept this as a paradox. If we allow circular definitions, we have such a paradox. To me it is merely a *reductio* of the stipulation. Why we cannot successfully make such a stipulation remains problematic, but the argument is a *reductio*, not a paradox. When Tarski used stipulation, it was to

---

[6] See Visser [2004, p. 153]. I use an equivalence here. Visser stipulates 'l' has the same meaning as '$\sim l$'.

say c abbreviates 'the sentence ...' and it was a fact that the sentence.... happened to be 'The sentence ... is not true'. I cannot see any reason to accept either the stipulation $q = \{x: x \notin q\}$ or the premise of Visser's purported Liar.

Secondly, there is the intuition that despite formal similarities the issues arise from the concepts concerned, i.e. the extension of each of the predicates and relations concerned: 'is true', 'is true of' or 'satisfies', and 'is a member of'. So, the Liar and satisfaction paradoxes are issues to be addressed by a theory of truth and more generally semantics; whereas Russell's is a problem which set theory can deal with independently. There is a bewildering variety of theories of truth which address the Liar independently on this basis. Fortunately, there are just two different conceptions of *set*: Cantor's conception of a set as a collection of objects and Frege's conception of a set as the extension of a predicate. Associated with these conceptions are a limited number of analyses of the pathology of sets. There are two types of associated problem in naive set theory. Problems with non-well-foundedness and problems with *impredicativity*. In the case of q, it purports to be the extension of a predicate. However, the set q is non-well-founded, as are Type-1 putative sets in general. In practice, non-well-founded sets are ruled out by the free variable constraint on abstraction and the axiom of Foundation. Russell's set and others are not avoided this way. In Russell's set-theory, there is no such set or class, it is *impredicative* to consider whether sets are members of themselves or not.

Thirdly, and associated with the above comments about Russell's concept of impredicativity, it is tempting to think of Type-2 paradoxes as *adding* something into the RA of the predicate. In the case of Grelling's, once one is introduced to the idea that predicates can be divided this way, one can consider whether Grelling's predicate is true of itself. In the case of Russell's set it is perhaps clearer. Having recognised an extension for 'is not a member of itself', one can ask if the extension itself of that predicate satisfies that predicate, that is, whether the extension contains the predicate itself. We do not think of Type-1s as adding anything. The Liar sentence is *prima facie* in the range of the truth-predicate. Russell's set however is not *prima facie* in the RA of its predicate, because the existence of the set is dependent on whether Russell's set is itself in RA of its predicate. Totalities initially seem to have little to do with the predicate case though, but if one thinks in terms of extensional semantics, a predicate cannot be completely interpreted unless its extension can be fixed.

Fourthly, there is the tempting but incorrect impression that I may have given in my presentation of the paradoxes that the Type-1 paradoxes necessarily contain terms

or even individual constants naming sentences that cannot be eliminated by canonical names containing no such non-canonical names for sentences. This impression encounters too many counter-examples. Quantificational variants like the Epimenides, for example, do not need to use non-canonical names. Furthermore, techniques like arithmetization enable us to form canonical names of expressions containing no non-canonical names. An identity can be proven using arithmetization and used to derive a Liar paradox. In chapter 5, I will give an example.

Fifthly, the Type-2 paradoxes may be proven without an identity premise, as I exemplified when I gave a one line proof of Grelling's paradox.

Let me set aside these intuitions for now, and introduce some working definitions for Type-1 and Type-2 paradoxes:

> In Type-1 *paradoxes* and *hypodoxes*, there is *prima facie* some thing in the RA of the predicate for which there is no way of determining whether the thing is in the extension of the predicate or not.

For the Liar, the truth of a suitable identity may be a contingency, or it may even be a theorem. Given the premise, then there is something to which the truth predicate should naively apply, either truly or falsely; but either assumption results in a contradiction.

Contrary to the expectations of some, it is not because the truth predicate itself is in its own range of applicability. The Liar is a sentence. The truth-predicate does not have a problem with being applied to itself, strictly speaking; it has a problem being applied to Liar-like sentences. *Prima facie* these are in the RA of the truth predicate. (When Liar sentences are theoretically excised from the RA of the truth predicate, there may be revenge or strengthened Liar problems.)

Similarly, for the Unsatisfied paradox, the truth of a suitable identity may be a contingency, or it may even be a theorem (which I will prove in the next chapter). Given the premise, then the predicate should naively apply to anything in the domain, either truly or falsely; but either assumption for anything in the domain leads to a contradiction.

One may stipulate the identity for a set q and use the same formal argument towards a contradiction; but it is not a paradox, because there was no reason to think the premise was true. By contrast with the Unsatisfied predicate, the putative set q fails a constraint on circular definition, or, in any case, its identity statement is provably false. One has an associated paradoxical argument, while the other has a *reductio*. For

there are independent reasons for thinking the identity premise of the Unsatisfied predicate is true. My favourite predicate just happens to be 'does not satisfy my favourite predicate'; and that identity is true independently of its consequences, because 'does not satisfy my favourite predicate' is a well-formed predicate and I favour it. Although I may think that my favourite set is the set of all things that are not members of my favourite set, I can only so favour it if it is a well-formed set. A correct semantic criterion for Type-1 paradoxes and hypodoxes needs to be sensitive to this contrast.

Here is a working or strawman definition of type2 paradoxes and hypodoxes.

> In Type-2 paradoxes and hypodoxes, there is no way of determining whether the predicate truly or falsely applies to the predicate itself or some function of the predicate.

Liars come with a reason for being in the RA of truth. Russell's set, may or may not be in the range of applicability for its own condition of membership. There are analogies which motivate us to think it might be; it may naively be thought to be one of the types of things that is in the RA of the membership predicate. There is pre-theoretically, prior to finding the paradox, a lack of any decisive reason why it should not be, except an alternate conception of membership. It appears that some sets are or can be in the RA of their own membership condition; and this motivates considering whether Russell's set is in the RA of its membership condition.

In the simplest case, there is no independent reason that assures whether Grelling's predicate is in its own RA; but some predicates are and (pre-theoretically) there is a lack of any reason why Grelling's predicate should not be in its own RA. The considerations that motivate naively considering Russell's set and Grelling's predicate as being in the RA of their own respective predicates are, although not conclusive, nevertheless effective. The arguments that ensue are represented in the tables in the previous section as the naive forms of these paradoxes. I am not suggesting that the solution is simply that these sets and predicates are not in those respective RAs.

There is a problem about how Grelling's predicate could conceptually not be in its own range of applicability. If the predicate *r* is not in its own RA, then it naively still seems that there is an extension for *r*, and it does not include *r*; so, *r* is not true of itself, therefore *r* is in its own RA. A similar line of *naive* thinking applies to Russell's paradox as well. The problem is similar to the strengthened Liar. If *w* were not in the RA of 'is not a member of itself', then it seems *w* would have an extension, and as *w* is

not in that extension, then *w* would not be a member of itself and therefore is in the RA of 'is not a member of itself'.

A naively simple *function* on a predicate, $\mathcal{P}$, is the set forming function $\{x:\mathcal{P}x\}$. If Russell's set is not in the RA of 'is not a member of itself', then it is clearly a fallacy to conclude that it is not a member of itself. For if it is not in the RA of that predicate, then it is a category mistake to say it is or is not a member of itself. There is a residual issue however. Even if *w* is not in the RA of 'is not a member of itself', naively, one still thinks there is a collection of sets that are not members of themselves, and it makes sense to ask whether *this* collection is a member of itself or not. The issue seems to be how *w* could not be in the RA of the membership predicate. This residual issue parallels the residual issue for a similar approach to Grelling's paradox.

So, in Type-2 cases, it seems more natural to question whether the contentious predicate, property or set itself is in the domain to be divided. Type-2 cases are more like the archetypal Russellian *vicious circles*, wherein a set *formed* by a division is itself a member of the domain to be divided. As is well-known, the referent of 'the tallest member of the basket ball team' is defined with respect to a totality of which he or she is a member, but this expression is unproblematic in its applications; because the tallest member of the basket ball team is someone, for which there is independent reason assuring us he or she is in the RA of 'is a member of the basket ball team'. In the case of Russell's set and Grelling's predicate there are no decisive reasons that settle whether the set or the predicate are in the RAs of their respective predicates.

Among the satisfaction paradoxes there are two types of problems associated with the two types of paradox. Firstly, there are problems over objects that have an independent reason to be in the range of applicability of truth or satisfaction. Secondly, there are problems over predicates themselves for which there is no independent reason why they themselves or the results of functions on them should not be in the RA of satisfaction or membership.

To put this another way, in Type-1 paradoxes, whether the sentence, predicate or set is in the domain is not problematic. The Liar sentence is grammatically well-formed and *prima facie* among the things to be determined true or false. The Unsatisfied predicate is grammatically well-formed. *Even if the Unsatisfied predicate is not in its own domain of meaningful application, the paradox can still be generated by taking any object that does lie within the predicate's range of meaningful applicability and assuming that object does or does not satisfy the predicate!* Its paradoxicality does not depend on whether it is in its own range of applicability.

4.4.2 Proof-theoretic Criteria

In any case, the two types of paradox are distinguishable proof-theoretically.

The derivations of Type-2 paradoxes are limited in the assumptions that can be used, generally to assuming they are reflexive or not. By contrast, the Unsatisfied predicate is paradoxical whatever it is applied to; nevertheless, this does not seem to lead directly to a formal distinction

For Type-2 paradoxes, the identity premises that I have been using are an unnecessary convenience. For example, the one-line proof below of Russell's paradox from no premises. It is an instance of the abstraction schema using an abstract canonical name:

$$\{x: x \notin x\} \in \{x: x \notin x\} \text{ iff } \{x: x \notin x\} \notin \{x: x \notin x\}$$

The examples I have given of Type-1 paradoxes use a non-canonical referring expression; so, an identity premise (or something to establish co-reference) is required. The Eubulidean Liar and the Unsatisfied paradox each require an identity premise (or effective equivalent) in addition to their schemata. It seems, as I have presented them, that they require non-canonical names. However, let us briefly consider three plausible counter-examples. Firstly, 'This sentence is false,' does not use an arbitrary name, but an indexical expression. The reference of the indexical must be given in context in order to know whether the sentence refers to itself or another. Even so, the reference of 'this sentence' cannot be effectively obtained from the expression 'this sentence' itself. So, co-reference must be established formally by an identity or something equivalent. Secondly, quantification can instantiate to canonical names for the problematic sentence, and, as in the Epimenides for example, give rise to paradox under certain circumstances. The derivation of a paradox will still use an identity; but it can use canonical naming. This seems like a good counter-example. Thirdly, consider using arithmetization to form canonical names and use Gödel's diagonal lemma to derive the Liar. Thus, it appears we have no need of the identity premise, whose place is taken by a biconditional which is an instance of the diagonal lemma, and therefore a theorem rather than a premise. Arithmetization and diagonalization appear to replace our previous reliance on a premise with a theorem. It could be argued that is some sense arithmetization is a form of stipulation, as the system of Gödel numbering must be specified. Nevertheless, the more considered point I wish to make is that the proof of a Type-1 paradox uses an identity, whether it is a premise or theorem.

Type-2 paradoxes, in contrast, do not need an identity statement in their derivation. They can use canonical names (formed without stipulation or indexicality or an empirical identity) in such a way that the derivation can proceed simply by substituting into the relevant schema.[7]

Perhaps some other level of generality applies. Tritely, yes – just the disjunction of the two types will give us a general type. Formally, the Type-2 paradoxes seem to use a subset of the logical principles used by the Type-1 proofs. For the Type-1 proofs seem to require an identity and use Leibniz's law (=E). The identity need not be a premise, but is proven as a theorem in some cases. I will resume this formal investigation in Chapter 5, where it turns out that Type-2 paradoxes do require one other logical principle that is not necessarily required by Type-1s.

4.4.3 Leibniz's Law and Substitution of Identicals

Leibniz's law and the Substitution of Identicals will begin to play a significant part in subsequent discussion, and particularly in the next chapter. I use the following two inference rules inter-changeably, and may mean either in context. I also use '=E' for either rule ambiguously.

Let Leibniz's law (the indiscernibility of identicals) be represented:

$\alpha = \beta$

$@(\alpha//\beta)$ iff $@(\beta)$

where $@(\alpha//\beta)$ is just like $@(\beta)$ except one or more occurrences of β are replaced with α.

I take Leibniz's law as inter-changeable with substitution of identicals. So, I shall also use (and may mean):

$\alpha = \beta$

$\underline{@(\beta)}$

$@(\alpha//\beta)$

---

[7] So, depending on the definition of *diagonalization*, either it is not required in the derivation of Type-2 paradoxes or it is achieved in one and the same step by using an instance of the relevant schema, e.g. the Satisfaction schema. In the next Chapter I settle on the latter view based on Jacquette's [2004] definition of diagonalization.

4.5 Converses and some better-known examples of Satisfaction Paradoxes not generally known as Satisfaction Paradoxes

This section begins with a further analysis of the types of satisfaction paradox taking into account the converse of the Unsatisfied predicate. Then, in pursuit of our broader agenda, I show how the Property paradox, Grelling's heterological paradox and Quine's Liar relate to paradoxes of satisfaction. The latter was put forward by Quine as a more economical Liar, a *prima facie* counter-example to the proposed distinction. There are at least three uses of 'true' in various paradoxes concerning truth, as an operator: 'it is true that', as a predicate: 'is true', and as a relation: 'is true of'. The converse of the truth relation is the satisfaction relation. Paradoxes of the truth relation are paradoxes of satisfaction. Historically, these have been associated with Grelling's paradox. Nevertheless, first we shall discuss the converse of the 'Unsatisfied predicate', which I used in my new 'Unsatisfied paradox'.

4.5.1 The Converse of the Unsatisfied Predicate

We may form the converse of the Unsatisfied predicate either by using the converse relation 'is true of' or by swapping the order of the term and the variable in the relation. Here is a converse:

	Paradoxical derivation using a Converse of the Unsatisfied Predicate:	
1	$p_2 = \langle\sim(p_2 \text{ Sat } x)\rangle$	Premise
2	$p_2 \text{ Sat } \langle\sim(p_2 \text{ Sat } x)\rangle \text{ iff } \sim(p_2 \text{ Sat } p_2)$	1, Satisfaction schema
3	$p_2 \text{ Sat } p_2 \text{ iff } \sim(p_2 \text{ Sat } p_2)$	2, Identity

Note that:

$$\sim(p_2 \text{ Sat } x) \equiv \sim(x \text{ is true of } p_2)$$

Notice also that the assumption that any arbitrary thing, t, satisfies p_2 does not lead to paradox, only its self-application. The converse of the Unsatisfied predicate is interesting because it is not so interesting as its converse. Since the Unsatisfied predicate is paradoxical whatever we assume to satisfy it, it follows by an analogous proof that the converse predicate will be paradoxical when we consider whether it satisfies itself. The converse, however, is not generally paradoxical.

From the above, we easily see that replacing 'p_2' in the right hand side of the identity with the variable 'x' will give the same result by analogous derivation. In fact

such a translation gives the Unsatisfied Grelling. However, in subsequent sections it will be clear that the simple mechanical replacement of a constant by a variable to obtain a reflexive expression is inadequate for mapping circular expressions from Type-1 to Type-2. We can clearly characterise circular and infinite expressions for Type-1 and Type-2. However, for quantified expressions, while the Type-1 cases are of some interest because there appears to be a Type-1 paradox set-theoretic paradox involving a quantified expression, there may not be distinct Type-2 expressions. These further issues will be analysed when they arise in carrying out a detailed mapping in section 6, but because of them the simple mapping is disconfirmed and a deeper analysis of the relationship is therefore required.

The converse of Unsatisfied predicate is paradoxical in a reflexive case, whether or not it satisfies itself, but not generally. It is a self-referential expression; but its paradoxical case is a case of self application. So, the converse of the Unsatisfied predicate is a caution to distinguishing referential from reflexive paradoxes.

Nevertheless, the converse of the Unsatisfied predicate meets the formal criterion for being a Type-1 paradox, as it requires an identity or equivalent to derive a paradox. In any case, the converse of the Unsatisfied predicate is certainly not equivalent to the Grelling predicate. Substitution instances of '~(x Sat x)' are a *reflexive subclass* of substitution instances of '~(x Sat y)'. Substitution instances of '~(x Sat *p*)', where '*p*' names that predicate itself, are a different *subclass*. The unit class of '~(*p* Sat *p*)' is the reflexive subclass of substitution instances of '~(x Sat *p*)'. These two *subclasses* overlap in just one case, '~(*p* Sat *p*)' – but this is not the "'heterological' is heterological" case. That latter case is the instance of '~(x Sat x)' where '~(x Sat x)' or its abbreviation or name is substituted for 'x' (and 'p' obviously is not an abbreviation or name for '~(x Sat x)').

4.5.2 The Property Paradox is Grelling-like

This one originates with Russell [Sainsbury 1995, p. 108]. Here is Sainsbury's version:

> Most properties are not true of themselves. For example, the property of being a man is not true of itself, since that property lacks the property of being a man; but the property of being a non-man is true of itself, since the property of being a non-man has the property of being a non-man. The pattern of reasoning used in the Class paradox [i.e. Russell's] would lead to the conclusion:

> The property of *being not true of itself* is true of itself if and only if it is not true of itself.
>
> There is at least a surface similarity between this contradiction and the contradiction that L_2 [the Strengthened Liar sentence] truly predicates truth of itself if and only if it does not. Where the Property contradiction uses the notion of *not true of*, a relation that may hold between a property and something else (perhaps also a property), the Liar contradiction uses the notion of *not true*, a property that a sentence or statement may possess.
>
> [Sainsbury 1995, p. 127]

The analogy Sainsbury draws breaks down because the property paradox is analogous to a Type-2 satisfaction paradox. Take angle brackets to name properties. Then a converse of the Satisfaction schema, an *is-true-of schema*, has the name of a property on one side of its biconditional, but uses a predicate on the other side. So, the Property paradox is analogous with the converse of the *Unsatisfied Grelling's Paradox*:

	The Untrue Grelling's Paradox:	
1	$r_2 = \langle$x is not true of x$\rangle$	Premise
2	$\langle$x is not true of x$\rangle$ is true of r_2 .iff. r_2 is not true of r_2	1, Is-true-of schema
3	r_2 is true of $r_2 \equiv$ r_2 is not true of r_2	2, Identity

The *Untrue Grelling* is very similar to Grelling's Heterological paradox. Grelling's original Heterological paradox actually involves adjectives. My simple formalism does not distinguish adjectives, but only has predicates and terms. My formal account uses the predicate:

$$r_2 = \langle\sim(\text{x is true of x})\rangle$$

This can be rendered less formally as 'x is not true of itself'.[8] If 'x' is replaced by a name of the adjective 'heterological', the result is:

---

[8] Quine [1962] abstracts this paradox from the heterological paradox.

'Heterological' is not true of itself.

In other words,

'Heterological' is not true of 'heterological'.

If the above were the case, however, given that 'heterological' applies to an adjective that is not true of itself, then:

'Heterological' is true of 'heterological'.

But given this latter result, the former follows again, and we have the dual *reductio* typical of an antinomy.

Grelling's heterological paradox is often grouped together with the Liar, as in Simmons [1993, p. 2]. However, my family break-up is between the Liar, Satisfaction and set-theoretic paradoxes. My basis for these families is roughly the schemata involved. The sense of this break-up is supported by the successful mapping of the classification within the Liar family into these other families, which I will demonstrate in subsequent sections. Some of my names belie this family break-up a bit. The *Untrue Grelling* is the converse of a satisfaction paradox and belongs to that family. It is a Type-2 Satisfaction paradox.

4.5.3 Quine's Liar

Quine proposed the following as a counter-example to the idea that a Liar paradox needed something more than the T-schema (whereas the Heterological paradox needed only an 'is true of' schema):

> " 'Yields a falsehood when appended to its own quotation' yields a falsehood when appended to its own quotation". This sentence specifies a string of nine words and says of this string that if you put it down twice, with quotation marks around the first of the two occurrences, the result is false. But that result is the very sentence that is doing the telling. The sentence is true if and only if it is false, and we have our antinomy.
>
> This is a genuine antinomy, on a par with the one about 'heterological', or 'false of self', or 'not true of self', being true of itself. But whereas that earlier one turned on 'true of', through the construct 'not true of self', this new one turns merely on 'true', through the construct 'falsehood', or 'statement not true'.

[Quine 1962, p. 7]

Following Quine's direction, consider:

> 'Yields a false sentence when appended to its own quotation' yields a false sentence when appended to its own quotation.

The above is true iff the below is true:

> " 'Yields a false sentence when appended to its own quotation' yields a false sentence when appended to its own quotation" is not true.

However, 'yields a false sentence' cannot stand as a grammatical unit on its own if it is to be applied to non-sentential expressions, the parsing must be to take 'yields a false sentence when appended to its own quotation' as a grammatical unit to be applied to non-sentential expressions. In considering the interpretation of this grammatical unit, let '⟨P⟩' be the name of a predicate and 't' be the name of some object, then:

> 'appending ⟨P⟩ to a name of *t* yields a true sentence' means ⟨tP⟩ is true.

So:

> " 'Yields a false sentence when appended to its own quotation' yields a false sentence when appended to its own quotation" means
> " 'yields a false sentence when appended to its own quotation' yields a false sentence when appended to its own quotation" is not true.

The above statement is analytically true. The two expressions are equivalent, thus:

> 'Yields a false sentence when appended to its own quotation' yields a false sentence when appended to its own quotation iff
> " 'yields a false sentence when appended to its own quotation' yields a false sentence when appended to its own quotation" is not true.

A contradiction follows from this equivalence together with an instance of the T-schema. Thus it appears we have a Liar that not only did not use a premise, but did not use an identity statement. However, analytic truths are not necessarily formal logical truths. This example exploited a simple grammatical rule that predicates appended to terms can yield meaningful sentences.

To represent all the logical commitments of this example, either something like an *append* function is required, or the meaning of the grammatical construction of appending something to itself can be represented by the truth relation.

In fact, in the first case, because the logical grammar uses a different order of term and predicate, it is easier to represent this with a norm function.[9] The norm function takes a predicate as an argument and yields a sentence composed of the predicate appended with the predicates own canonical name:

norm(⟨$\mathcal{P}$⟩) = ⟨$\mathcal{P}$ ⟨$\mathcal{P}$⟩⟩

1. norm('~T(norm)') = '~T(norm('~T(norm)')'	definition
2. T⟨~T(norm('~T(norm)')⟩ iff ~T(norm('~T(norm)')	T-schema
3. T(norm('~T(norm)') iff ~T(norm('~T(norm)')	2, 1 =E

Thus Quine's Liar, interpreted this way, does avoid the need for a premise, but still formally uses an identity statement.

In the second case, if Quine's analytic statement is derived from a relation, then its grammatical use of the reflexive pronoun can be represented.

> 'appending ⟨P⟩ to a name of *t* yields a true sentence' means '⟨P⟩ is true of *t*'.

So:

> 'Yields a statement not true when appended to its own quotation' yields a statement not true when appended to its own quotation.

means

> 'yields a statement not true when appended to its own quotation' is not true of 'yields a statement not true when appended to its own quotation'.

Quine's mere 'true' is really an 'is true of' in disguise. Quine's argument is analogous to the following.

> 'Is not true of itself' is not true of itself .

means:

> 'Is not true of itself' is not true of 'is not true of itself',

---

[9] Norms were devised following Quine by Smullyan [1957]. See also Smullyan [1994, p.4-7].

and so, by the is-true-of schema, is equivalent to:

" 'Is not true of itself' is not true of itself" is not true.

From this equivalence and an instance of the T-schema:

" 'Is not true of itself" is not true of itself" is true .iff. " 'Is not true of itself" is not true of itself" is not true.

Under this interpretation, Quine's paradox requires no premise, but it did tacitly use 'is true of', not merely 'true'. Quine *is* right about this paradox using the T-schema. Consider again, the one line proof of the *Unsatisfied Grelling's paradox*, an instance of the Satisfaction schema:

⟨~x Sat x⟩ Sat ⟨~x Sat x⟩ iff ~(⟨~x Sat x⟩ Sat ⟨~x Sat x⟩)

from which, using the T-schema:

⟨~x Sat x⟩ Sat ⟨~x Sat x⟩> iff ~T⟨⟨~x Sat x⟩ Sat ⟨~x Sat x⟩⟩

This is analogous to Quine's result. The absence of a premise defining a name *turns* on 'is true of' in Quine's paradox and the instance of the Satisfaction schema using canonical names above, rather than the T-schema. However, the use of the T-schema here was inessential to deriving a contradiction.

Thus Quine's paradox, on the one hand as a Liar, is derived using a function and an identity statement, and, on the other hand, is a hybrid of Grelling's paradox; but relies on the proof of Grelling's paradox for proving a contradiction. Quine's paradox focuses attention to a distinction in formal proofs of the paradoxes, rather than simply the expression 'yields a falsehood when appended to its own quotation'.

4.6 Mapping

This section demonstrates the mapping of Liar paradoxes to Type-1 Satisfaction paradoxes, and then to Type-2 paradoxes, and to set-theoretic paradoxes. The distinction between Type-1 and Type-2 is exemplified in every case including quantified, circular and infinite versions.

In the previous chapter, I presented a classification of Liar paradoxes by slicing them into:

1. Self-referential Liar-like paradoxes and hypodoxes
2. Circular Liar-like generalisations of these.

3. Infinite Liar-like paradoxes and hypodoxes

Among the self-referential and circular, we have non-quantificational versions and versions involving quantified statements. I show how non-quantificational self-referential versions relate to quantificational versions. The self-referential are a subset of the circular – circles of one. The infinite Liars relate via lists. I made a *completeness* claim relating to sentences *a* through *h*, and *k* through *o* that I listed in a number of tables exemplifying all 12 semantic *possibilities* for self-referential unquantified sentences to be paradoxical or hypodoxical under some circumstances.

Among this set of sentences, the mapping of the Liar itself has been previously dealt with, and I now map the Unquantified Epimenides, the Self-referential Curry, and the ESP into paradoxes of satisfaction and membership. I then map quantificational sentences of this nature into satisfaction paradoxes. Then I map circular, and finally infinite variations in this way. Finally, I give a summary of the mapping of hypodoxes.

4.6.1 The Unquantified Epimenides

The table below arranges the paradoxes mapped from the Unquantified Epimenides as we did for the Eubulidean Liar in Section 2. The bottom left-hand space could be filled with a Type-1 Set-theoretic contradiction, but here and where these are not even naively acceptable as sets, we will not bother. Given any one of these Epimenidean identities, ~Q follows.

Type-1	Type-2
Unquantified Epimenides $b = \langle Tb \ \& \ Q \rangle$	
Unsatisfied Unquantified Epimenides $p_b = \langle \sim(x \text{ sat } p_b) \ \& \ Q \rangle$	Unsatisfied Epimenidean Grelling $r_b = \langle \sim(x \text{ sat } x) \ \& \ Q \rangle$
	An Unquantified Epimenides set $w_b = \{x: x \notin x \ \& \ Q\}$

4.6.2 Curry's

Here is the mapping for variations on Curry's paradox. (There are other variations of implicational versions of Curry's paradox, which I consider to be beyond the scope of the present exercise.) Given any one of these identities, Q follows.

Type-1	Type-2
Self-referential Curry $c = \langle \sim Tc \vee Q \rangle$	
Unsatisfied Curry $p_c = \langle \sim(x \text{ sat } p_c) \vee Q \rangle >$	Unsatisfied Curried Grelling $r_c = \langle \sim(x \text{ sat } x) \vee Q \rangle$
	Curry's paradox $w_c = \{x: x \notin x \vee Q\}$

4.6.3 The ESP Set

The Eldridge-Smith paradox (the ESP) is relatively new, and I have given the proofs in the table below, although I must say they follow the same pattern. Nevertheless, the use of the identities as premises in the Type 2 cases is an abbreviational convenience, one could begin the type 2 proofs in an unabbreviated form of their second line as an instance of the satisfaction schema.

Type-1	Type-2
The ESP	
$d = \langle \sim Td \leftrightarrow Q \rangle$ $T\langle \sim Td \leftrightarrow Q \rangle$ iff. $\sim Td \leftrightarrow Q$ Td iff. $\sim Td \leftrightarrow Q$ $Td \rightarrow (\sim Td \leftrightarrow Q)$ $Td \rightarrow \sim Q$ $\sim Td \rightarrow \sim(\sim Td \leftrightarrow Q)$ $\sim Td \rightarrow (\sim Td \leftrightarrow \sim Q)$ $\sim Td \rightarrow \sim Q$ $\sim Q$	
The Unsatisfied Self-referential ESP	Unsatisfied Reflexive ESP

$p_d = \langle\sim(x \text{ sat } p_d) \leftrightarrow Q\rangle$ t sat $\langle\sim(x \text{ sat } p_d) \leftrightarrow Q\rangle$ iff $[\sim(t \text{ sat } p_d) \leftrightarrow Q]$ $(t \text{ sat } p_d)$ iff $[\sim(t \text{ sat } p_d) \leftrightarrow Q]$ $(t \text{ sat } p_d) \rightarrow (\sim(t \text{ sat } p_d) \leftrightarrow Q)$ $(t \text{ sat } p_d) \rightarrow \sim Q$ $\sim(t \text{ sat } p_d) \rightarrow \sim(\sim(t \text{ sat } p_d) \leftrightarrow Q)$ $\sim(t \text{ sat } p_d) \rightarrow (\sim(t \text{ sat } p_d) \leftrightarrow \sim Q)$ $\sim(t \text{ sat } p_d) \rightarrow \sim Q$ $\sim Q$	$r_d = \langle\sim(x \text{ sat } x) \leftrightarrow Q\rangle$ $(r_d \text{ sat } r_d)$ iff $[\sim(r_d \text{ sat } r_d) \leftrightarrow Q]$ $(r_d \text{ sat } r_d) \rightarrow [\sim(r_d \text{ sat } r_d) \leftrightarrow Q]$ $(r_d \text{ sat } r_d) \rightarrow \sim Q$ $\sim(r_d \text{ sat } r_d) \rightarrow \sim[\sim(r_d \text{ sat } r_d) \leftrightarrow Q]$ $\sim(r_d \text{ sat } r_d) \rightarrow [\sim(r_d \text{ sat } r_d) \leftrightarrow \sim Q]$ $\sim(r_d \text{ sat } r_d) \rightarrow \sim Q$ $\sim Q$
	The ESP set
	$w_d = \{x: x \notin x \leftrightarrow Q\}$ $w_d \in w_d$ iff. $w_d \notin w_d \leftrightarrow Q$ abstraction $w_d \in w_d \rightarrow . w_d \notin w_d \leftrightarrow Q$ $w_d \in w_d \rightarrow \sim Q$ $w_d \notin w_d \rightarrow \sim(w_d \notin w_d \leftrightarrow Q)$ $w_d \notin w_d \rightarrow \sim(w_d \notin w_d \leftrightarrow \sim Q)$ $w_d \notin w_d \rightarrow \sim Q)$ $\sim Q$

4.6.4 Quantification

Famously, as the story is told in Chapter 1, Epimenides the Cretan said, 'Cretans are always liars', and gave his name to a related paradox, *The Epimenides*. I have given the derivation of the paradox using quantification in Chapter 1, and discussed how the Church conundrum is derived from an Epimenides-like statement, and that if the contingent circumstances support the extra premise a contradiction results. In Chapter 3, I discussed how the Epimenides relates to the Eubulidean Liar and other members of the Liar family.

Here is the corresponding Type-1 paradox in terms of satisfaction, where 'C' stands for some condition (or relation):

	Type-1 Unsatisfied Epimenides	
1	$p_s = \langle(\forall y)(Cy \supset \sim(x \text{ Sat } y)\rangle$[10]	premise
2	Cp_s	premise
3	t sat p_s iff $(\forall y)(Cy \supset \sim(t \text{ Sat } y))$	Sat schema, 1 Id
4	| t Sat p_s	assumption
5	| $(\forall y)(Cy \supset \sim(t \text{ Sat } y)$	3, 4 SL
6	| $C\, p_s \supset \sim(t \text{ Sat } p_s)$	5 ∀E
7	| $\sim(t \text{ Sat } p_s)$	6, 2 MP
8	$\sim(t \text{ Sat } p_s)$	4-7 ~I. (Nothing satisfies p!)
9	$\sim(\forall y)(Cy \supset \sim(t \text{ Sat } y)$	8, Sat schema, equivalences
10	$(\exists y)(Cy \,\&\, t \text{ Sat } y)$	QN, definition, DeM: Conc. 2

Here, the identity statement is used as an abbreviational convenience. In Chapter 1, I have shown that the Epimenides has minimal dependence on the use of such an identity statement. I will discuss quantification in relation to its dependencies again at the end of Chapter 5.

In the above, rather than have p_s explicitly self-referential, we use quantification that includes p_s in its domain.[11] Condition C looks strange. The conclusion says that for any individual, there is something that the individual satisfies which satisfies the condition. This is analogous to Church's conundrum. If we add a premise that nothing other than p_s satisfies the condition C, we have a contradiction. (Note that "converses" of p_s such as p_3 where $p_3 = \langle(\forall y)(Cy \supset \sim(y \text{ Sat } x)\rangle$ are also paradoxical but the derivation must make a reflexive assumption like 'p_3 Sat p_3'. This does not make these converses Type-2 paradoxes though, because they require contingent premises.)

There is a Type-1 Set-theoretic paradox in this case, involving the *Epimenides Set*. Because the quantifier binds the 'y' variable, it satisfies the restriction on the abstraction schema:

---

[10] Alternatively, one could have a relation Cxy, such that $p_s = \langle(\forall y)(Cxy \supset \sim(x \text{ Sat } y))\rangle$ and for the second premise: Ctp_2. Another alternative premise set is: $p_s = \langle(\forall y)(Cx \supset \sim(x \text{ Sat } y))\rangle$ and Ct, but the above better preserves the analogy with the Epimenides as it binds the variable in the antecedent.

[11] Curiously, without explicit universal quantification, the open sentence is formally just a special case of Curry's; whereas semantically it is more intuitive to relate the Epimenides to conjunctions like the Unquantified Epimenides, $b = \langle\sim Tb \,\&\, Q\rangle$.

	Type-1 Set-theoretic Epimenides	
1	$q_s = \{x: (\forall y)(Cy \supset x \notin y)\}$	premise
2	Cq_s	premise
3	$t \in q_s$ iff $(\forall y)(Cy \supset t \notin y)$	Abstraction schema, 1 Id
4	| $t \in q_s$	assumption
5	| $(\forall y)(Cy \supset t \notin y)$	3 MP
6	| $C\, q_s \rightarrow t \notin q_s$	$\forall$E
7	| $t \notin q_s$	MP
8	$t \notin q_s$	~I, Conc. 1: Nothing satisfies q_2
9	$t \notin \{x: (\forall y)(Cy \supset x \notin y)\}$	Identity
10	$\sim(\forall y)(Cy \rightarrow t \notin y)$	9, sat schema, equivalences
11	$(\exists y)(Cy \,\&\, t \in y)$	QN, definition, DeM: Conc. 2

Unlike q, the identity at line 1 does not fail as a definition of a set for circularity, at least it is not ruled out by Boolos' free variable constraint. It is hardly news that if t satisfies some condition, it must be a member of some set. However, if it were also given that t is not a member of any other set satisfying the condition C, such as being referred to from a certain page of a certain book, this would be paradoxical.

Even though the definition of q_s is not ruled out by Boolos' free variable constraint, this argument is not paradoxical. *Prima facie*, it is a reductio of the disjunction of line 1 and 2. Line 2 could conceivably be a true premise, but as nothing independently justifies line 1 (unlike the Liar's premise), then it is a *reductio* of the identity given in line 1.

In contrast however, Quantified Type-2 paradoxes appear to be insignificantly distinguishable from unquantified Type-2 paradoxes. Indeed,

$$r_s = \langle(\forall y)(Cy \supset \sim(x \text{ sat } x)\rangle$$

has no interesting derivations that make it worth distinguishing as another category of paradox. The derivation below demonstrates this.

	Quantified Unsatisfied Grelling	
1	$r_s = \langle(\forall y)(Cy \supset \sim(x \text{ sat } x)\rangle$	premise

2	Cr_s[12]	premise
3	r_s sat r_s iff ⟨(∀y)(Cy ⊃ ~(g sat g)⟩	Sat schema
4	| r_s sat r_s	assumption
5	| (∀y)(Cy ⊃ ~(r_s sat r_s)	MP
6	| Cr_s ⊃ ~(r_s sat r_s)	∀E
7	| ~(r_s sat r_s)	MP
8	~(r_s sat r_s)	~I
9	~(∀y)(Cy ⊃ ~(r_s sat r_s)	Identity
10	(∃y)(Cy & r_s sat r_s)	QN, defn, DM
11	r_s sat r_s	∃E, &E (N.B. y is bound in r_s)
12	#	11, 8 &I: Contradiction (#)

I note, for completeness, that a version of Russell's paradox could be derived along analogous lines, but is also not really worth distinguishing as a variation of Russell's paradox.

Here is the summary table for this set of paradoxes:

Type-1	Type-2
The Epimenides with identity	
1 | s = ⟨∀x (Dx ⊃ ~Tx)⟩ 2 | Ds or D⟨∀x (Dx ⊃ ~Tx)⟩ for both premises … 11 | ∃x (Dx & Tx) … And with an additional premise: 12 | ∀x (Dx & x ≠ s .⊃ ~Tx) … 15 | #	
The Unsatisfied Epimenides	The reflexive unsatisfied Epimenides
C⟨∀y(Cy ⊃ ~(x Sat y)⟩)	C⟨∀y(Cy ⊃ ~(x sat x)⟩

[12] Here, Ct would do just as well. It would perhaps be better to use a relation 'Cxy' in the identity and then have Crt as the second premise.

... $\forall x\ (Cx\ \&\ x \neq \langle \forall y(Cy \supset \sim(x\ Sat\ y))\rangle) .\supset$ $\sim(t\ Sat\ x))$	... #
The Epimenides set	The Epimenidean Russell set
$q_s = \{x: \forall y(Cy \supset x \notin y)\}$ Cq_s ... $t \notin q_s$ $(\exists y)(Cy\ \&\ t \in y)$ ... $\forall x\ (Cx\ \&\ x \neq q_s .\supset t \notin x)$	$w_s = \{x: \forall y(Cy \supset x \notin x)\}$ Cw_s ... $w_s \notin w_s$ ... $w_s \in w_s$ #

4.6.5 Proper Circularity

Circular Liars are often presented as lists of two or more sentences like this:

(1) Sentence (2) is true.

(2) Sentence (1) is not true.

but can also be formed using lists of conjuncts, like this:

The next conjunct is false and the previous conjunct is true.

Here is a summary of some examples of circular paradoxes:

Type-1	Type-2
Circular Liar $(1) = \langle \sim T(2)\rangle$ $(2) = \langle T(1)\rangle$	
Circular Unsatisfied paradox $p_1 = \langle \sim(x\ sat\ p_2)\rangle$ $p_2 = \langle (x\ sat\ p_1)\rangle$	Circular Unsatisfied Grelling $r_3 = \langle \forall y \sim(x\ sat\ y\ \&\ y\ sat\ x)\rangle$
	Circular Russell's $w = \{x: \forall y \sim(x \in y\ \&\ y \in x)\}$

With respect to the Circular Russell, this and versions with longer chains are found in Quine [1940 / 1951, pp. 128-129].

We cannot form a circular Type-2 satisfaction paradox by using a list of predicates in the way we can form a Type-1 satisfaction paradox. If we attempt to do so by particularizing 'y' instead of using a variable, we will end up with a Type-1 paradox instead. Furthermore, 'y' as a variable must be bound in both instances by the same quantifier to assure that 'y' will be instantiated with the same thing. We could, along the lines I mentioned, use conjunctions to represent both types as follows:

> Type-1: x does not satisfy the next conjunct and x satisfies the previous conjunct.
>
> Type-2: For all y, it is not the case that x satisfies y and y satisfies x.

The two types of paradoxes can be extended to *n* conjuncts, and a contradiction is obtainable in all cases. A difference worth noting is that in a circle of three statements where two deny the truth of subsequent members, as below, we get what Sorensen calls a *no-no* paradox, such that (1) can be true, and (2) and (3) false; or (1) could be false, and (2) and (3) true.[13]

$= \langle \sim T(2) \rangle$

$= \langle T(3) \rangle$

$= \langle \sim T(1) \rangle$

In contrast, when we expand the list of conjuncts in the Circular Russell as below (or its unsatisfied correlate), it is not clear whether there is any corresponding effect.

$$w = \{x: \forall y \, \forall z \sim(x \in y \,\&\, y \in z \,\&\, z \in x)\}$$

This may have other ramifications, but for current purposes it further confirms the distinctness of Type-1 and Type-2 paradoxes.

4.6.6 Infinite Versions

I illustrate the mapping of infinite versions with Yablo's paradox. (The same can be done for Sorensen's Queue Paradox.) For Yablo's paradox, we consider an infinite sequence of statements, the n-th member of which is:

---

[13] I argue in the previous chapter that "*no-no* paradoxes" are really hypodoxical duals of paradoxical combinations.

$Y_n = \langle$For all $k > n$, S_k is untrue$\rangle$.

As well as the T-schema, we also use the non-controversial schemata:

$T\langle\forall k{>}n\ {\sim}TY_k\rangle \rightarrow \forall k{>}n{+}1\ {\sim}TY_k$

$T\langle\forall k{>}n\ {\sim}TY_k\rangle \rightarrow {\sim}TY_{n+m}$

Goldstein [2000] has presented a set-theoretic version of Yablo's paradox and the Queue paradox respectively. There is a technical difficulty: it relies on an infinite sequence of sets, for which Goldstein [2000, p. 68] gives the *n*-th member by means of a statement as:

$Z_n = \langle\forall x\ (x \in C_n \leftrightarrow$ for all $k > n$, $(x \notin C_k)\rangle$

Each sentence, Z_n, in the sequence is taken as specifying a class, C_n.[14] In the above, the 'k' variable is bound. For any Z_n, the sets it refers to, or quantifies over, form an infinite sequence of some sort, as do Yablo's infinite series of statements. Whether these sets exist is independent of whether or not there are sentences about them. So, we may dispense with Goldstein's sequence of Z-sentences and simply consider the sets:

$Q_n = \{x: \forall k{>}n\ (x \notin Q_k)\}$, for all $n \in \mathbf{N}$

As for our criteria for whether this is a Type-1 or Type-2 paradox: Q_k is not a mere abbreviation. We cannot do without the identity premise(s) and use instances of the abstraction schema free of names like 'Q_k'. Furthermore, I can show that for any thing, *t*, it is both a member of some class in this sequence and is not a member of any class in the sequence. Thus Goldstein's example meets the criteria for being more like the Liar than Russell's. I put it that Goldstein's is a Type-1 set-theoretic paradox. It meets all the criteria. I provide the analogous Type-1 satisfaction paradox in the table

---

[14] Goldstein [2000, p. 72] suggests the Infinite Liars have a circular model as the tail of a finite sequence can point back to the beginning and thereby model an infinite number of subsequent like statements. Goldstein wants to argue that the Infinite versions are essentially circular, contrary to Sorensen's [1998] defence of the non-circularity of Yabloesque paradoxes. However, making a member of the above sequence of class abstraction statements point back to the beginning will not work. It would not then specify an infinite number of classes. Furthermore, classes are not abstracted based on whether or not *a statement* of that abstraction *exists*. A circular sequence of Liar-like statements does not have the same difficulties. However, if the loop has an even number of statements, it will be a "no-no" paradox, not a Liar paradox. In an even numbered loop, every second statement can be false while each alternate one is true; however, we have no basis for distinguishing which is which. As a model then it must have an odd number of statements, yet no such restriction applies to the infinite sequence it purports to model.

below, which illustrates the complete analogy in the proofs of these paradoxes with Yablo's.

Type-1		
Yablo's	The Infinite Unsatisfied Yabloesque Paradox	Goldstein's Infinite Set-theoretic Liar paradox
$Y_1 = \langle \forall k>1 \sim TY_k \rangle$ ... $Y_n = \langle \forall k>n \sim TY_k \rangle$ ...	$P_1 = \langle \forall k>1 \sim(x \text{ Sat } Y_k) \rangle$ ... $P_n = \langle \forall k>n \sim(x \text{ Sat } Yk) \rangle$ ...	$Q_1 = \{x: \forall k>1 (x \notin Q_k)\}$ ... $Q_n = \{x: \forall k>n (x \notin Q_k)\}$ ...
$\mid TY_1$ $\mid \forall k>1 \sim TY_k$ $\mid \sim TY_2$ $\mid \forall k>2 \sim TY_k$ $\mid TY_2$ (TI) $\sim TY_1$ $\sim\forall k>1 \sim TY_k$ (TE) $\mid TY_a$ $\mid \forall k>a \sim TY_k$ $\mid \sim TY_{a+1}$ $\mid \forall k>a+1 \sim TY_k$ $\mid TY_{a+1}$ $\sim TY_a$ $\forall k>1 \sim TY_k$ #	$\mid t \text{ Sat } P_1$ $\mid \forall k>1 \sim(t \text{ Sat } P_k)$ $\mid \sim(t \text{ Sat } P_2)$ $\mid \forall k>2 \sim(t \text{ Sat } P_k)$ $\mid t \text{ Sat } P_2$ $\sim(t \text{ Sat } P_1)$ $\sim\forall k>1 \sim(t \text{ Sat } P_k)$ $\mid t \text{ Sat } P_a$ $\mid \forall k>a \sim(t \text{ Sat } P_k)$ $\mid \sim(t \text{ Sat } P_{a+1})$ $\mid \forall k>a+1 \sim(t \text{ Sat } P_k)$ $\mid t \text{ Sat } P_{a+1}$ $\sim(t \text{ Sat } P_a)$ $\forall k>1 \sim(t \text{ Sat } P_k)$ #	$\mid t \in Q_1$ $\mid \forall k>1 (t \notin Q_k)$ $\mid t \notin Q_2$ $\mid \forall k>2\ t \notin Q_k$ $\mid t \in Q_2$ $t \notin Q_1$ $\sim\forall k>1\ t \notin Q_k$ $\mid t \in Q_a$ $\mid \forall k>a, t \notin Q_k$ $\mid t \notin Q_{a+1}$ $\mid \forall k>a+1, t \notin Q_k$ $\mid t \in Q_{a+1}$ $t \notin Q_a$ $\forall k>1\ t \notin C_k$ #

Here is my own Type-2 set-theoretic version corresponding to Yablo's. It is the Yabloesque version of Russell's paradox. In this case, w_i contains a set x just if there is no infinite sequence of sets the first of which is a member of x and each subsequent set in the sequence is a member of its predecessor. Think of an *infinitely descending sequence of sets* as ordered by membership. Then w_i is the set of all sets that are not the first members of an infinitely descending sequence of sets. w_i cannot be a member of itself, because it would be the fist member of an infinitely descending sequence of sets, in particular the sequence of sets (w_i, w_i, w_i, ...). Because w_i is not a member of

itself, there is no infinitely descending sequence of sets of which w_i is the first member, which means w_i is a member of itself – contradiction. We might represent w_i as:

$w_i = \{x: \forall s \sim(\forall k \in \mathbf{N}\ s_{k+1} \in s_k \ \&\ s_1 = x)\}$

where '$\forall s$' is taken to quantify over infinite sequences of sets; and 's_1' refers to the first member of that sequence, 's_k' refers to the *k*th member of that sequence, etc.

Alternatively, where '$\forall y_{k\in\mathbf{N}}$' stands for '$\forall y_1 \ldots \forall y_k \ldots$', here is a proof wherein a contradiction arises between lines 8 and 15 (hereing 'w' is abbreviates 'w_i'):

	My Type-2 Infinite Set-Theoretic Paradox		
1	$w = \{x: \forall y_{k\in\mathbf{N}} \sim(y_1\in x \ \&\ y_2\in y_1 \ \&\ y_3\in y_2 \ \&\ \ldots y_n\in y_{n-1} \ \&\ldots)\}$		Premise
2	$w \in \{x: \forall y_{k\in\mathbf{N}} \sim(y_1\in x \ \&\ y_2\in y_1 \ \&\ y_3\in y_2 \ \&\ \ldots y_n\in y_{n-1} \ \&\ldots)\}$ iff $\forall y_{k\in\mathbf{N}} \sim(y_1\in w \ \&\ y_2\in y_1 \ \&\ y_3\in y_2 \ \&\ \ldots y_n\in y_{n-1} \ \&\ldots)$		Abstraction
3	$w \in w$ iff $\forall y_{k\in\mathbf{N}} \sim(y_1\in w \ \&\ y_2\in y_1 \ \&\ y_3\in y_2 \ \&\ \ldots y_n\in y_{n-1} \ \&\ldots)$		1,2 Identity
4		$w \in w$	Assumption
5		$\forall y_{k\in\mathbf{N}} \sim(y_1\in w \ \&\ y_2\in y_1 \ \&\ y_3\in y_2 \ \&\ \ldots y_n\in y_{n-1} \ \&\ldots)$	3,4 MP
6		$\sim(w\in w \ \&\ w\in w \ \&\ w\in w \ \&\ \ldots w\in w \ \&\ldots)$	5 $\forall$E w/y_1, w/y_2, $\ldots w/y_n$, ...
7		$w\notin w$	6 Simplification (of an infinite conjunction)
8	$w\notin w$		4-7 ~I
9		$a_1\in w \ \&\ a_2\in a_1 \ \&\ a_3\in a_2 \ \&\ \ldots a_n\in a_{n-1} \ \&\ldots$	Assumption
10		$a_1\in w$	9 &E
11		$\forall y_{k\in\mathbf{N}} \sim(y_1\in a_1 \ \&\ y_2\in y_1 \ \&\ y_3\in y_2 \ \&\ \ldots y_n\in y_{n-1} \ \&\ldots)$	Abstraction, 1 Id, 10 MP
12		$\sim(a_2\in a_1 \ \&\ a_3\in a_2 \ \&\ \ldots a_n\in a_{n-1} \ \&\ldots)$	11 $\forall$E a_2/y_1, a_3/y_2, $\ldots a/y_n$, ...
13	$\sim(a_1\in w \ \&\ a_2\in a_1 \ \&\ a_3\in a_2 \ \&\ \ldots a_n\in a_{n-1} \ \&\ldots)$		9-12 ~I
14	$\forall y_{k\in\mathbf{N}} \sim(y_1\in w \ \&\ y_2\in y_1 \ \&\ y_3\in y_2 \ \&\ \ldots y_n\in y_{n-1} \ \&\ldots)$		13 $\forall$I y_1/a_1, y_2/a_2, $\ldots y_n/a_n$, ...
15	$w\in w$		3,14 MP

There is an analogous Type-2 satisfaction paradox. And here is a summary of our collection of Yabloesque paradoxes:

Type-1	Type-2
Yablo's Paradox $Y_1 = \langle \forall k>1 \sim TY_k \rangle$ ... $Y_n = \langle \forall k>n \sim TY_k \rangle$	
Unsatisfied Yablo $P_1 = \langle \forall k>1 \sim(x \text{ Sat } Y_k) \rangle$ ... $P_n = \langle \forall k>n \sim(x \text{ Sat } Yk) \rangle$ ...	My Infinitely Unsatisfied Paradox $r_i = \langle \forall p \sim(\forall k \in \mathbf{N}\ p_{k+1} \text{ Sat } p_k \ \&\ p_1 = x) \rangle$ where '$\forall p$' quantifies over infinite sequences of predicates, the first member of which is p_1, the *k*th member is p_k, etc.
Goldstein's Infinite Set-theoretic Paradox $Q_1 = \{x: \forall k>1\ (x \notin Q_k)\}$ ... $Q_n = \{x: \forall k>n\ (x \notin Q_k)\}$ ...	My Infinite Set-theoretic paradox $w_i = \{x: \forall s \sim(\forall k \in \mathbf{N}\ s_{k+1} \in s_k \ \&\ s_1 = x)\}$ where '$\forall s$' quantifies over infinite sequences of sets, the first of which is s_1, the *k*th is s_k, etc.

4.6.7 Mapping Truth-teller-like Hypodoxes

Finally, I briefly map *hypodoxes* (Truth-tellers). The Satisfied Truth-teller, p_e, is such that anything may or may not satisfy it:

$$p_e = \langle x \text{ Sat } p_e \rangle$$

Whereas the Satisfied Grelling below is only problematic with respect to itself, the Satisfied Grelling, like its 'autological' version, typically is indeterminate about whether it satisfies itself or not.

$$r_e = \langle (x \text{ Sat } x) \rangle$$

The corresponding hypodox in set-theory is the set of all self-membered sets:

1. $w_e = \{x: x \in x\}$
2. $w_e \in w_e$ iff $w_e \in w_e$ 1. abstraction

For which there seems no fact of the matter as to whether it is a member of itself or not – it seems it can be either.

Each mapping can be carried out this way, but for brevity I provide only a few more set-theoretic examples. The Unquantified Epimenidean Hypodoxical set is:

$$w_f = \{x: x \in x \vee \sim Q\}$$

From which, by abstraction:

$$w_f \in w_f \text{ iff. } w_f \in w_f \vee \sim Q$$

Now, if Q is the case, there is no fact of the matter whether w_f is a member of itself.

The Curried Truth-teller corresponds to:

$$w_g = \{x: x \in x \ \& \sim Q\}$$

From which, by abstraction:

$$w_g \in w_g \text{ iff. } w_g \in w_g \ \& \sim Q$$

Now, if ~Q is the case, there is no fact of the matter whether w_g is self-membered.

And in the case of the ESP, I note that its hypodoxical set-theoretic version still entails Q (as does its semantic 'truth-teller', which was sentence *h* in the summary table of section 5).

1. $w_h = \{x: x \in x \equiv Q\}$
2. $w_h \in w_h \text{ iff. } w_h \in w_h \equiv Q$ 1. abstraction
3. $w_h \in w_h \rightarrow (w_h \in w_h \equiv Q) :\&. (w_h \in w_h \equiv Q) \rightarrow w_h \in w_h$
4. $w_h \in w_h \rightarrow Q$
5. $w_h \notin w_h \rightarrow \sim(w_h \in w_h \equiv Q)$
6. $w_h \notin w_h \rightarrow (w_h \notin w_h \equiv Q)$
7. $w_h \notin w_h \rightarrow Q$
8. Q

So, if Q is the case, there is no fact of the matter whether w_h is self-membered on not.

In any case, I think I have provided sufficient mappings to substantiate that the distinction between Type-1 and Type-2 paradoxes and hypodoxes can be implemented for all relatives of the Liar paradox.

4.7 Conclusion

The project to classify paradoxes has been progressed. We began by identifying two types. Ironically, we achieve some reduction in the paradoxes that challenge us in this way, a plausible division of labour or a subset on which to focus further analysis. We can take a classification of the Liar family and map it to the first branch of the satisfaction family. Members of the first branch generally have counter-parts in the second branch. Members of the second branch can be mapped into set-theoretic paradoxes. (There are still other set-theoretic paradoxes, such as Cantor's paradox, not directly map to in this process.)

At this level of detail, the distinction between Type-1 and Type-2 paradoxes are not a basis for Ramsey's distinction. The line is shifted to a distinction between two types of satisfaction paradox. Furthermore, some Type-1 paradoxes circumvent circularity restrictions in naive set-theory, meaning there are some set-theoretic paradoxes on both sides of the distinction.

Priest's [2000, p. 123] Principle of Uniform Solution: 'same kind of paradox: same kind of solution' is a significant philosophical claim. It would be trite to respond 'Two types of paradox: possibly two problems: possibly two solutions required'. There is a mapping from Liars to one type of Satisfaction paradox, then these map intuitively to the second type, which have analogous set-theoretic paradoxes. Therefore, there ought to be a composite mapping from Liars into the set-theoretic paradoxes. However, the mapping between Type-1 and Type-2 satisfaction paradoxes remains intuitive at this stage.

The split between different branches of satisfaction paradoxes appears to align with folk beliefs in the distinction between referential and reflexive paradoxes, but complexities arise, such as Quine's Liar, which warrant further reflection on the relation between paradoxical arguments and expressions. Nevertheless, a distinction seems to exist even in this complex case between two types of paradox, based on their formal proofs, which will be the primary focus of the next chapter.