Journal of Philosophical Logic 33 (2004): 27-88

A Contextual-Hierarchical Approach to Truth and the Liar Paradox*

Michael Glanzberg
Department of Linguistics and Philosophy
MIT
Cambridge, MA 02139
USA
glanzber@mit.edu

June 3, 2003

*Thanks to Chris Gauker, Warren Goldfarb, Volker Halbach, Ned Hall, Richard Heck, Olafur Jonsson, Charles Parsons, Carol Voeller, Ralph Wedgwood, and an anonymous referee for helpful comments and discussions.

Abstract

This paper presents an approach to truth and the Liar paradox which combines elements of context dependence and hierarchy. This approach is developed formally, using the techniques of model theory in admissible sets. Special attention is paid to showing how starting with some ideas about context drawn from linguistics and philosophy of language, we can see the Liar sentence to be context dependent. Once this context dependence is properly understood, it is argued, a hierarchical structure emerges which is neither *ad hoc* nor unnatural.

Your bait of falsehood takes this carp of truth:
And thus do we of wisdom and of reach,
With windlasses, and with assays of bias,
By indirections find directions out:
Hamlet II.i.68–71

It is a perennial idea in the study of the Liar paradox, from Tarski [56] onwards, that its solution requires some kind of *hierarchy*. More recent, but also quite common, is the idea that the paradox is to be solved by some sort of appeal to *context dependence*. Both ideas face a fundamental objection: they appear to be *ad hoc* solutions, unmotivated by our prior understanding of the concepts involved, and incompatible with the applications we make of them. Hierarchical solutions appear to require multiple truth predicates, which fragments what is intuitively one concept into many. Moreover, they impose restrictions on the application of the truth predicate that are incompatible with many of its ordinary and apparently unproblematic uses (as seminal work of Kripke [32] made abundantly clear). The idea of context dependence appears to fare no better. It simply does not appear that the sentences which generate the Liar paradox contain context-dependent elements, nor do we have any clear idea what aspect of context might be involved in the paradox.

In this essay, I shall attempt to show that a combination of the ideas of context dependence and hierarchy can provide a *non-ad hoc*, well-motivated, and intuitively attractive approach to the Liar paradox. I shall argue that we really do have good reason, independent of the paradox itself, to see the Liar sentence as context dependent. I shall also argue that once we have a proper understanding of the nature of the context dependence involved, we find that a hierarchical structure of semantic relations, including the truth predicate, *emerges*. I shall thus argue that neither context dependence nor hierarchy are *ad hoc* responses to the paradox, but genuine features of the concepts involved. The hierarchy that so emerges is more liberal than the sort suggested by Tarski (and so avoids the objections from Kripke), but is able to resolve the paradoxes in an essentially hierarchical way.

Much of the work in this paper is directed towards identifying and explaining the context dependence that is at work in the Liar. Once we understand it well enough, we can move on to explain how it leads to hierarchical structures. The paper is divided into three parts. Part (1) will make heavy use of concepts from linguistics, particularly the linguistics of context dependence. It will identify the context-dependence phenomenon at work in the Liar, and sketch out an account of how it can have the kinds of effects that lead to a resolution of the paradox. Part (2) will pursue a mathematical model of the theory developed from empirical considerations in Part (1). Taken together, Parts (1) and (2) do the heavy lifting: they explain the basic phenomena that underly the Liar, and show why the apparently intractable paradox is not intractable after all. Part (3) focuses on the role of hierarchies. It returns to some of the details of the Liar that are discussed in the beginning of Part (1), and shows how they can be resolved by further developing the the apparatus provided by Part (2). This will show how the hierarchies that emerge from context dependence can be used resolve the problem of the Liar.

1 Context

In this part, I shall offer an informal explanation of the context-dependent nature of the Liar. In the first Section (1.1), I shall argue that a highly unusual sort of context dependence—which I shall dub 'extraordinary context dependence'—is at work. In the next Section (1.2), I shall sketch an account of this phenomenon. I shall argue that though extraordinary, the context dependence involved does find motivation in the linguistics of context dependence. I shall hence argue it is not an *ad hoc* posit. I shall then outline an explanation of how it can have the drastic effects we see with the Liar paradox. In the next Part (2), I shall provide a detailed formal development of the claims I motivate and outline here.

1.1 Context Dependence and the Liar

Why would anyone think there is any context dependence involved in the Liar paradox? The basic reason is what is known as the *Strengthened Liar paradox*. Consider a Liar sentence: for the moment, let us take a sentence l_s which says l_s is not true. ('l_s' for 'sentential'.) We can, informally, reason as follows. First, we observe that l_s is true just in case it is not true. Hence, it is tempting to say, l_s must lack a truth value altogether. Gap theorists may find this appealing, but it is no safe haven. For from this it follows that l_s is not true. Observe, our last conclusion was just l_s itself. This is a correct conclusion, hence true, contra our prior claim. Indeed, we can infer that l_s is true after all. We are back in paradox. This is called the Strengthened Liar paradox, as it shows that our initial attempt to avoid the Liar by declaring the Liar sentence to lack a truth value ultimately fails to avoid paradox.

The strengthened paradox stems from two conclusions made in succession:

1. l_s is not true.

2. l_s is true.

How could one avoid such a clear contradiction?[1] One strategy is to observe that in fact these are two distinct utterances. If we could argue that there is a shift in context between them, we could argue that there is only an appearance of contradiction. A sentence is used in one utterance, and the negation of that sentence is used in another. But because there is a difference in context, this does not mean there is in fact a contradiction. Compare two other utterances, made as we point to cars passing by us:

1. That is loud.

2. That is not loud.

There need be no contradiction here, if we are pointing to different cars as we say each of these. This would be a difference in context. The hope is that something similar might get us out of the Strengthened Liar.

Demonstrative expressions like 'that', as well as indexicals like 'I', 'here', etc., are well-known to induce context dependence. Unfortunately, there is little intuitive reason to suppose that any such expression appears in l_s. Is there then any hope of context-based resolution of the Strengthened Liar?

1.1.1 Linguistic Context Dependence

I believe there is. Before investigating the Liar sentence in particular, let me begin by noting that the linguistic phenomenon of context dependence goes well beyond the classes of indexicals and demonstratives. We see context dependence, for instance, in gradable adjectives like 'tall'. Roughly, tall seems to mean something like 'tall for an X', where X is a comparison class. This can be provided by context, so saying 'John is tall' can roughly mean tall for a basketball player, or tall for a jockey. We also see context dependence in some complex constructions, such as genitives. 'John's team' can mean, depending on the context, the team John likes, the team John hates, the team John owns, etc.

[1] I believe the term 'Strengthened Liar' is due to van Fraassen [60]. It is often presented as a separate problem from the Liar, or as a kind of 'revenge paradox', but I take it to be the fundamental issue, as do, for instance, Burge [3], Gaifman [10, 11], and Parsons [43]. Gaifman [12] notes that the term 'Strengthened Liar' has been applied to a number of different phenomena. The version I discuss starting in Section (1.1.3) below, and take to be the fundamental issue, is in essence what he calls the 'unable-to-say' paradox.

One sort of context dependence will be of particular importance for what follows: the phenomenon of contextual quantifier domain restriction. Consider the claim 'Everything was destroyed by the fire'. In spite of there being no overt restriction on the quantifier 'everything', an utterance of this normally only claims that some restricted domain of things were destroyed—those in the house, for instance. The domain is set by context.[2]

But even with this expanded list of context-dependence phenomena from natural language, we are left stymied. It seems evident that the Liar sentence l_s contains no indexicals, no demonstratives, no genitives, no gradable adjectives, and no quantifiers. It seems that there is just no context dependence in the Liar, whether or not that would provide a happy solution to the puzzle. And, we might add, there is little reason to see an appropriate context shift between the utterances (1) and (2) in the Strengthened Liar. The is nothing in it comparable to the introduction of a salient object, as there was in passing from (1) to (2) in the 'That is loud' case. What hope have we of using context dependence at all?

A number of proposals have been made for why the Liar sentence is context dependent after all, in spite of appearances to the contrary. With Parsons [43], I shall ultimately propose that there is a quantifier domain restriction phenomenon at work, though an extraordinary one. A more common idea [3, 30, 48, 49] is to suppose that the truth predicate itself contains a hidden indexical component.[3] Let me briefly note that I do not think this is a promising option. It is a commonly voiced objection to it that we simply do not intuitively see such an indexical element in our ordinary truth predicate, expressed by the ordinary term 'true'. I believe this line of argument can be bolstered. If there were such a hidden indexical, it would behave as other implicit parameters do. In the case of a gradable adjective, for instance, we can see the hidden comparison class at work when we *bind* the hidden variable, as in:

[2]The literature on quantifier domain restriction is quite large. Two useful starting places are Stanley and Szabó [55] and von Fintel [62].

[3]Koons [30] and Simmons [48, 49] build upon the work of Burge [3], but make some significant departures. How far their commitments to an index on the truth predicate go is not entirely easy to judge. For instance, the object language in some of the formal discussion of Simmons [48] does not contain an indexed truth predicate, though indexed sentences are evaluated and indexed truth predicates are introduced. In later work [49], Simmons does appear to more thoroughly rely upon an indexed truth predicate.

In locating context dependence in the Liar sentence in a quantifier domain restriction, I take my motivations from natural-language semantics. Another approach that avoids positing an index on the truth predicate is that of Gaifman [10, 11, 12], which takes its motivations from the semantics of programming languages.

Most species S have members that are small for S.

We see no such behavior with the truth predicate.[4]

1.1.2 Expression and Quantifier Domains

Where then should we look for context dependence in the Liar? Perhaps we should return to basics. A general description of the phenomenon of context dependence can be given if we help ourself to some familiar apparatus. Roughly, context-dependent sentences 'say' different things upon different occasions of use. Let us introduce *propositions* as the contents expressed by utterances in contexts. A *sentence* is then context dependent if it expresses different *propositions* in different *contexts*. As we saw a moment ago, there is a large range of constituents of sentences that can induce context dependence.

With this in mind, let us return to the truth predicate itself. Above, I applied it to the *sentence* l_s, which seemed fine for the application at hand. But in the presence of context dependence at large, it is well-know that this cannot be exactly right; for the same sentence could be both true and false (depending on the context)! Given we have the apparatus of propositions available, they are the obvious candidates to be truth bearers. Now, the question of what the fundamental bearers of truth are, and whether propositions have the right nature to be them, is a much-debated philosophical puzzle. I shall say as little about this matter here as I can get away with. I shall assume propositions somehow encapsulate truth conditions, as anything that provides the content of an utterance should. As all that concerns us here is truth and related concepts, I shall therefore work as if propositions are simply sets of truth conditions, though I do not want to take a stand on whether this is a simplifying assumption, or a correct analysis. But once we have propositions as truth conditions, then it is entirely natural to see them as bearers of truth.[5]

[4] The use of this sort of test for implicit parameters is discussed by Stanley [54] and Stanley and Szabó [55]. Stanley mentions the problem for Burge's theory, though I take the discussion to follow to show it is not a problem for all context-dependence approaches. For some alternative diagnostics, see Larson [33] and Ludlow [36].

[5] The claim that propositions are unstructured sets of possible worlds is defended by Stalnaker [52]. Arguments that propositions must be structured abound. See, for instance, Soames [50]. Very little in my discussion hangs on the nature of propositions. For the most part, interpreted sentence tokens, or utterances, would work just as well, so long as we take these things to be within the ranges of appropriate quantifiers. Some ways to develop proposals similar to mine without the use of propositions are discussed in Parsons [43].

This does not yet make much difference. If there is no context dependence, then the difference between applying truth to the sentence l_s, or the proposition it invariably expresses, is of no consequence. But it will help to reveal a source of additional context dependence. Actually, the real source is in a related notion. If we apply truth to propositions, then for most applications of the truth predicate we will make, we will have to work with a proposition presented via a sentence that expresses it. Though we might sometimes manage to name a proposition in some more or less direct way, we usually wind up saying something like 'the proposition expressed by sentence s is true', or more informally, 'what she said when he said s is true'.

It is this relation which introduces an additional source of context dependence. Most predications of truth, including those relevant to the Liar, work through the expression relation. As such, they contain a tacit existential quantifier. To say 'the proposition expressed by s in context c is true' is to say

$$\exists p(\text{Expresses}(s, p, c) \wedge \text{True}(p)).$$

Like all quantifiers, this one *can* have its domain contextually restricted. Suppose we have a specific domain of proposition in mind: say, those that some particular person might say. Then we can use 'Every proposition is true' to mean that every proposition in the domain is true, just as we can use 'Everything was destroyed by the fire' to mean everything in a particular house was destroyed.

This reveals an element that can support context dependence in the apparatus of truth predication. It does not show that any particular use of it is context dependent. (Indeed, the example I just gave is somewhat contrived.) Nor does it show that a contextual domain restriction is at work in the Liar. But it opens the door to just this idea.

1.1.3 Ordinary and Extraordinary Context Dependence

We now have a quantifier over propositions at work in predications of truth, and so in the Liar sentence. We thus have an opening for context dependence. The problem we face is that it is still hard to see any restriction on the domain of this quantifier. Unlike my example above, we cannot simply appeal to speakers' intentions to explain how its domain is restricted. If anything, we would expect speakers intentions to make this quantifier *un*restricted. Where then lies the context dependence? Moreover, it is not yet clear why such a domain restriction would matter. As we noted, the Liar sentence still contains no indexicals, or any of the other devices of context dependence we canvassed. If this implies that there is a unique proposition expressed by the Liar sentence, contextual domain restriction would become virtually irrelevant.

To settle on some terminology, I shall say that the argument just given shows that there is no *ordinary context dependence* involved in the Liar. There are none of the common effects of indexicals or the ordinary sorts of contextual domain restrictions we have considered, etc. I shall argue that nonetheless there is context dependence at work, which must be an *extraordinary context dependence*.

To see what extraordinary context dependence might be, it will be helpful to regiment the Liar sentence more precisely. As I mentioned, predications of truth will go via the expression relation. Normally, this is a three-place relation between a sentence, a proposition, and a context. However, the accepted lack of ordinary context dependence in the Liar will allow us to make a simplification and work with a two-place expression relation $\mathsf{Exp}(s, p)$ between a sentence and a proposition. For the example at hand, the only potential source of context dependence is the propositional quantifier, which will be explicitly represented in the logical form of the sentence. So we will have no need to make use of the third argument of the expression relation, and can safely ignore it.[6]

To present the logical form of the Liar sentence, we will also need a truth predicate of propositions, which I shall write $\mathsf{Tr}(p)$. Now, we may consider the Liar sentence. It is a sentence which 'says' of itself that it is not true. This is a sentence l which is:

$$\neg\exists p(\mathsf{Exp}(\ulcorner l \urcorner, p) \wedge \mathsf{Tr}(p)).$$[7]

This formulation displays the lack of ordinary context dependence we saw above. There is no potentially context-dependent element except the quantifier $\exists p$. It is natural to presume this quantifier has no ordinary restriction either. Under its most natural interpretation, we do not expect speakers to mean there is no true proposition expressed by l in some specific domain—we expect them to mean there is no proposition, period. Hence, this formulation presumes that there is no ordinary context dependence at work in the Liar.

[6] I have discussed this issue at greater length in [14].

[7] To minimize the use of quotation marks, when I introduce syntactic items I often put them in sans serif and their interpretations in *italics*. I shall assume we have some device of sentence naming, indicated by ⌜corners⌝. In Part (2), this will be a Gödel coding. Where it can be done without confusion, I often suppress quotation marks. I shall hence often write s instead of $\ulcorner s \urcorner$.

I take the the Strengthened Liar reasoning to show that there must be some context dependence in the Liar nonetheless, and thus that there must be extraordinary context dependence. To reproduce the Strengthened Liar, I shall rely on three natural assumptions about Tr and Exp. One is a basic principle about truth:

$$\text{(T-Exp)} \qquad \mathsf{Exp}(\ulcorner\sigma\urcorner, p) \to (\mathsf{Tr}(p) \leftrightarrow \sigma).$$

Another reflects the idea that an utterance expresses a unique proposition. As we are assuming no ordinary context dependence, this is:

$$\text{(U-Exp)} \qquad (\mathsf{Exp}(s,p) \wedge \mathsf{Exp}(s,q)) \to p = q.$$

The final one is an instance of a law of identity which marks that truth is an extensional property of propositions:

$$\text{(T-Id)} \qquad p = q \to (\mathsf{Tr}(p) \leftrightarrow \mathsf{Tr}(q)).$$

We may now produce a version of the Liar paradox. First, suppose l expresses a proposition, say q. We then have $\mathsf{Exp}(\ulcorner l\urcorner, q)$. We can argue that q is true just in case it is not true.

1. Suppose $\mathsf{Exp}(\ulcorner l\urcorner, q)$.
2. Suppose $\mathsf{Tr}(q)$.
   (a) By (T-Exp), $\neg\exists p(\mathsf{Exp}(\ulcorner l\urcorner, p) \wedge \mathsf{Tr}(p))$.
   (b) Hence, $\neg\mathsf{Tr}(q)$.
3. Conversely, suppose $\neg\mathsf{Tr}(q)$.
   (a) Then again by (T-Exp), $\exists p(\mathsf{Exp}(\ulcorner l\urcorner, p) \wedge \mathsf{Tr}(p))$.
   (b) By (U-Exp), $p = q$ and $\mathsf{Tr}(p)$.
   (c) Hence, by (T-Id), $\mathsf{Tr}(q)$.

We thus have $\mathsf{Tr}(q) \leftrightarrow \neg\mathsf{Tr}(q)$. This is not yet a paradox, of course, for it is reached under the assumption $\mathsf{Exp}(\ulcorner l\urcorner, q)$. What we have is a proof of:

$$\neg\exists p \mathsf{Exp}(\ulcorner l\urcorner, p).$$

But a paradox does follow. From this conclusion, by logic, we may conclude:

$$\neg\exists p(\mathsf{Exp}(\ulcorner l\urcorner, p) \wedge \mathsf{Tr}(p)).$$

Observe, this is just l. Now, we have proved l, so it had better be true. But to be true is to express a true proposition. So we have in fact shown:

$$\exists p \mathsf{Exp}(\ulcorner l \urcorner, p).$$

This does indeed put us in paradox.

Let us call this reasoning the *Liar inference*. In the course of it, we make two crucial assertions:

(A) $$\neg \exists p (\mathsf{Exp}(\ulcorner l \urcorner, p).$$

(B) $$l.$$

As both are the results of sound proofs, they must both be true. But the truth of the first requires that there be no proposition for l to express, while the truth of the second requires that there be one. Hence the paradox.

I maintain we have to see a context shift between (A) and (B) affecting l. We have a proof, based on solid principles, so its conclusions had better be correct. If there is no context shift, then we have a genuine contradiction. As we have already ruled out any ordinary context dependence in l, this shows there must be some extraordinary context dependence.

The only locus for context dependence, ordinary or extraordinary, is the domain of the quantifier $\exists p$. We can now see what context dependence must provide. This domain must *expand* between (A) and (B). If this is occurs, both claims can be correct. The first claim (A) may be true, as relative to the domain of propositions available from its context, there is no proposition for l to express. The second claim (B) may express a true proposition, as relative to the expanded domain of propositions available from its context, there is one.

This is indeed an extraordinary demand. It is clear that in the course of the proof, speakers intentions and other ordinary sources of context dependence do not restrict the domain of $\exists p$. The conclusion at (A) is that there is no such p, not that there is no such p in some specifically restricted domain. We have a proof, carried out in full generality. Even so, I shall argue this does not preclude extraordinary context dependence. In particular, I shall argue that we have good reason to see the *maximal domain* over which $\exists p$ can range in a given context as apt to expand as context shifts, and that this induces just the kind of extraordinary context dependence we need.

1.1.4 The Expansion Problem

I have labeled the problem of explaining this sort of extraordinary context dependence the *expansion problem*. Solving the expansion problem confronts us with two distinct tasks.

First, we have to explain how we can have an expansion of the maximal domain over which the quantifier $\exists p$ can range in a given context. As I mentioned above, in a discussion of truth and the Liar paradox, it is harmless to identify propositions with sets of truth conditions: something like sets of possible worlds, or whatever else we use to model individual truth-supporting circumstances. For the maximal domain of $\exists p$ to expand is precisely for the background domain of truth conditions from which propositions are formed to expand. The contextual expansion of the maximal domain of the propositional quantifier reduces to the expansion of the background domain of truth conditions.

But second, we must not lose sight of the problem of explaining just what the shift in context between (A) and (B) is, and how that particular change can lead to this sort of domain expansion. Indeed, as I have discussed at length elsewhere [14], it is not at all obvious that there is any shift in context between (A) and (B). After all, these are merely two points in an *inference*, marked by no changes in the immediate environment of the speakers of the kind that were so evident in the 'That is loud' example at the beginning of Section (1.1).

1.1.5 Constraints on a Solution

Before confronting these tasks, let me pause to make one comment on the way I shall approach them. There is a relatively direct way to address the issues I have raised, very closely related to the situation-based proposal of Barwise and Etchemendy [2]. I shall not take the direct route. But to motivate the route I shall take, and to explain why I think it offers a viable solution, I shall begin by discussing the direct route, and what I think is attractive and unattractive about it.

In a spirit similar to that of Barwise and Etchemendy, let us start with the idea that instead of being a possible world, an individual truth condition is a *situation*. The situation theory of Barwise and Etchemendy is quite specific, but for this discussion, let us take a situation to be a *partial* and *structured* representation of some part of a possible world. The situation of a being F can be the structure $\langle a, F, 1\rangle$. The situation of a not being F would then be $\langle a, F, 0\rangle$.

Because situations are structured and partial, we can easily make sense of the expansion of a background domain of situations available in a context. Suppose that a context somehow hands us the materials—the properties and objects—from which situations may be constructed. Suppose that in our initial context, we have only non-semantic properties: some range of properties excluding the expression relation and the truth predicate. If in a subsequent context, these semantic relations become available, we can use them to construct new situations, distinct from any of the situations available in the previous context. We will have new situations of the form $\langle \ldots, \mathsf{Exp}, \ldots \rangle$ and $\langle \ldots, \mathsf{Tr}, \ldots \rangle$. As situations are structured entities, we can be certain that these situations are distinct from any available in the previous context. Expansion problem solved.[8]

As we will see, I shall ultimately argue in favor of a view that has many features in common with this idea. But as a starting point, I find it is unappealing. It appears to simply insist that truth conditions or situations with truth predicates tacked on are different from those without. No explanation of this has been given, and indeed, the way in which truth-supporting circumstances are supposed to *determine* what is true in them may well make us wonder if this is even a correct picture.[9] As Barwise and Etchemendy note, this proposal has a Tarskian air. And presented as I did, it may well appear just as *ad hoc* as Tarski's did. At the very least, it does not yet constitute an adequate rejoinder to the objection that context-dependent

[8]In detail, the proposal of Barwise and Etchemendy [2] is somewhat different. What I glossed as situations are closer to their states of affairs. Their situations are sets of states of affairs. A proposition is a pair of a situation and a situation type (built from states of affairs by closure under conjunction and disjunction). For instance, a proposition of the form $\{s; [\sigma]\}$ is true iff $\sigma \in s$. Relative to an initial situation s, they produce a Liar proposition $f_s = \{s; [Tr, f_s; 0]\}$. This is a proposition f_s which says that it does not hold in s. There is a sense in which this proposition cannot be expressed . In particular, the state of affairs $\langle Tr, f_s; 0 \rangle$ cannot be in s. (Actually, Barwise and Etchemendy say that the proposition is expressible, but give up on what they call the F-closure of s. But there is a core observation in common between these two points, and the details do not matter for our purposes here.) There is then a distinct situation $s' = s \cup \{\langle Tr, f_s; 0 \rangle\}$, and the proposition $\{s'; [Tr, f_s; 0]\}$ relative to this new situation—this new 'context'—is true.

We can bring this closer to the idea I have just sketched by noting that the truth-making situation for f_s contains the state of affairs $\langle Tr, f_s; 0 \rangle$. It is adding this explicitly semantic state of affairs to the initial situation that allows for the truth of the Liar proposition. This state of affairs cannot be in the initial situation. Hence, we start with some situation, and add a semantic state of affairs to it, resulting in a genuine expansion of the situation. This is a close analog to the kind of expansion I propose. (Barwise and Etchemendy discuss relations between their situation-based and a more traditional approach in [2, Ch. 11]. For a detailed match-up between the Barwise and Etchemendy framework and a Burgean framework of indexed truth predicates, see Koons [30].)

[9]This is often glossed as the supervenience of the semantic upon the non-semantic. (I find this issue less straightforward than it may appear. See my [15].) It makes the difference between situations like s and $s \cup \{\langle Tr, f_s; 0 \rangle\}$ hard to understand from a truth-conditional perspective, and makes it equally hard to understand just what (truth-conditional) content is given by a proposition represented as $\{s'; [Tr, f_s; 0]\}$.

and hierarchical approaches are intolerably *ad hoc*.

I do not have any quarrel with the use of situations *per se*, which has become increasingly common in linguistics.[10] But I am not a partisan of situation theory as a general framework for semantics, and have expressed some misgivings about offering them as a replacement for more standard notions of truth conditions. Hence, what I am asking for is a well-motivated theory which gets similar results, in a more conventional framework. Can we explain, starting from ideas of a more conventional nature from linguistics and philosophy of language, how the domain of truth conditions expands as context shifts?

In asking for this, I am asking as much as is possible to smoothly integrate what we say about the expansion problem, and extraordinary context dependence, into linguistic theory. Though I granted that extraordinary context dependence is extraordinary, it is still context dependence. To understand it, we need to explain how it fits into our understanding of the rest of how language works. No doubt to address a paradox as recalcitrant as the Liar, we will have to say something which goes beyond standard theorizing in linguistics and the philosophy of language. I doubt there would be a recognizable paradox if we did not. But to defend a solution to the paradox that is immune to the objection of being *ad hoc*, we need to build up as much as we can from more standard ideas.

1.2 Solving the Expansion Problem

My strategy for solving the expansion problem is thus to build up to it from some ideas from linguistic theory. These will provide an account of what sort of context shift takes place with in the Liar inference between (A) and (B) of Section (1.1.3). Some general ideas about truth conditions from the philosophy of language will then explain how this context shift leads to the expansion of the domain of truth conditions. Once these ideas are sketched in this section, I shall turn to a more formal model in Part (2).

1.2.1 Context in the Liar Inference

To explain extraordinary context dependence, as a species of context dependence, we need to start by looking a little more closely at context itself.

[10] A notable example is Kratzer [31], as well as von Fintel [62]. Other theories in semantics prefer events to situations, but either is usually offered as part of an overall truth-conditional framework, rather than as a replacement for truth conditions.

The basic idea that I shall make use of is that context provides a running record of information available at a given point in a discourse. This idea appears in a number of ways in recent thinking about context, but the one that will be important here is that context provides a running record of what is *salient*; particularly, what is *salient in a discourse* at a particular point.[11]

This idea has played a significant role in some recent work on (ordinary) context dependence. Let me briefly mention a couple of examples, to make the idea clearer. One important example comes from work on the referents of definite noun phrases, starting with Karttunen [27] and developed by Heim [23] and Kamp [26].[12] To use an example from Heim, suppose I say 'I broke a wine glass last night. It was expensive'. Here, the indefinite 'a glass' sets up a contextually salient *discourse referent*, which 'it' is then able to refer back to. The context provides the salient discourse referent, to which the definite noun phrase then refers. Note, it is the use of the indefinite in the discourse that makes a particular glass salient for the discourse. It would still be so, for instance, even if the conversation were taking place in a glass store.

The application of salience I shall make goes well beyond that of the referents of definites. But the literature already points to a much wider area of application. To mention one more example, consider the following dialogs (from Büring [5]):

1. **Q:** What did you buy on 59th Street?

   A: On 59th Street, I bought the shoes.

2. **Q:** Where did you buy the shoes?

   A: #On 59th Street, I bought the shoes.

These are naturally read with a rising pitch contour on the fronted prepositional phrase 'on 59th Street' and a falling one on 'the shoes'. In its normal reading, the second is infelicitous (hence the '#' marking). Why? Very, very roughly, because the context, fixed by the initial question, fails to make what happened on 59th Street appropriately salient in the discourse. The fronted constituent, marked by syntactically by fronting and prosodically by pitch

[11]The idea that context is a running record of information is central to the presupposition theory of context, articulated in Stalnaker [51], and appears as the notion of conversational record in Lewis [34]. Stalnaker's theory shows how the context set—the set of presupposed propositions—is updated as a discourse progresses. I have no quarrel with taking the context set to be part of context, but I have argued in [14] that it cannot help in accounting for the Liar, and so we must look to other aspects of context.

[12]For more recent references see [59].

accent, is marked as a topic: what the claim is about. Generally, these need to refer to entities that have the right sort of salience properties in the discourse. In this case, the topic-marked constituent is a prepositional phrase, taking us beyond the realm of definite noun phrases.

The role of topic in context has been pursued at great length, not only by Büring, but by such authors as Portner and Yabushita [44], Reinhart [45], Roberts [46] , Vallduví [58], van Kuppevelt [61], and von Fintel [62]. A related notion is basic to centering theory [19], applied to some linguistic issues related to topic by Ward [64].[13] Remarkably similar ideas have been applied to issues in natural language processing, as the focus spaces of Grosz [18] and the focusing structures of Grosz and Sidner [20].[14]

I have rushed over a rather large area, and I should note that there are important differences between the semantics of topic and of definite noun phrases. But I hope to have made clear, both from a couple of examples, and from reference to the linguistics and natural language processing literature, that taking context to include a running list of salient items is well-motivated. Moreover, I hope to have made equally clear that this needs to be taken quite widely: we need to keep track not only of salient discourse referents for noun phrases, but of a wide range of salient items of all sorts.

Given the richness of natural language—syntactic and semantic—it is not easy to give a linguistically accurate assessment of just what we need to keep track. But for our purposes, we may abstract away from some of this complexity. Thinking about the structure of formal languages, we should expect context to provide a running list of salient properties and individuals. As we expect to have quantifiers in our language, we should also include salient domains of quantification. Again simplifying somewhat, and prescinding from philosophical worries about the intensional nature of properties, this means that we may take this aspect of context to be given by a *structure* in the usual logician's sense. Context provides a structure

[13]The linguistics of topic is a messy and controversial subject, and there are important differences among the theories I have cited. Perhaps most important, some take topics to be questions under discussion rather than as salient items *per se*. I have argued that there are close connections between these ideas in my [16]. Elements of both appear in van Kuppevelt's work [61]. For some comparisons between these kinds of approaches, see McNally [39].

[14]I hasten to note some significant complications I am glossing over. The notion of focus space is somewhat more complicated than that of a list of salient entities. Focus spaces are intended to organize complex data structures. Grosz [18], in particular, starts with a 'semantic network' represented by a directed graph, and sees a focus space as a partition of it. Once one sees data structures as basic, for instance, one has the possibility of nesting focus spaces.

which reflects salient domains, individuals, and relations at a given point in a discourse. Call this structure the *salience structure*.

Salience structure is a component of context. I have in rather cursory fashion suggested there is good linguistic support for this claim. I have argued this more fully elsewhere [16].[15] Though the notion of salience structure abstracts from a good deal of linguistic complexity, I do wish to stress that the idea of salience structure *as a component of context* is linguistically well-motivated. This is important for my purposes here, as the linguistic motivation is entirely independent of any considerations of the Liar paradox. This will help to counter the objection that the contextual approach is *ad hoc* or unmotivated.

One of the crucial features of salience structure as an aspect of context is the way it changes as a discourse progresses. This is evident in the wine glass example above, for instance. At the beginning of the discourse, it is unlikely the pronoun 'it' could refer to a wine glass. It does after the indefinite 'a wine glass' makes it salient. To stress, the notion of salience here is salience at a point in a discourse: what we are generally talking *about* at that point. As a discourse progresses, this tends to change. Even in a discourse on a single topic, speakers move from subtopic to subtopic. One way to introduce a new subtopic is to introduce a new term into a discourse, in such a way as to mark its interpretation as salient for the discourse. Any time this is done, it induces a change in what is salient at that point in the discourse. In particular, it introduces a new salient item. The result is that the salience structure *expands*.[16]

This is the feature that makes salience structure the right aspect of context for addressing the Liar paradox. In the Liar inference of Section (1.1.3), two assertions (A) and (B) were identified, and I argued there must

[15] In [16] I call salience structure by the name 'topic structure'. My primary aim there is to show that the information involved in this aspect of context is information about the global structure of the discourse in which an utterance takes place. Hence, insofar as the structure is of salient items, it is salience for a discourse organized a particular way that is at issue, rather than salience in the immediate environment of an utterance. None of this conflicts with the applications of the notion I make here. The shift in terminology is to be as ecumenical as possible about the linguistic motivations for this aspect of context, and to mark idealization for purposes of application in a more formal setting.

[16] Of course, I am again abstracting away from a great deal of linguistic detail. I discuss at length how topics may change in discourse in [16]. I am further simplifying by ignoring the fact that in many discourses, salient items may be removed as well as added. It is not relevant to the application to the paradox I shall make.

Some computational linguists see networks of salient items (focus spaces) as kept on a stack. This is clear in Grosz and Sidner [20]. Walker [63] argues that a cache model is in order instead.

be a context shift between them. Attention to salience shows us what it is. The first context involves an assertion of $\neg\exists p \mathsf{Exp}(\ulcorner l \urcorner, p)$. Moreover, this assertion corresponds to the first point in the proof above where there are no undischarged premises. Hence, I suggest, it is the point where the relation *Exp* (which interprets Exp) is accepted as salient in the discourse.[17] But then, with the assertion (A), the salience structure is expanded to include *Exp*. The context thus shifts, by expanding its salience structure, so we have a genuine difference in context between (A) and (B), just as the paradox requires.

In what follows, I shall assume that a null context includes in its salience structure some minimal information about the language being spoken, including some syntactic information. Hence, again abstracting from a great deal of linguistic complication, we may take a null context to have a salience structure suitable for doing Gödel coding. We may assume such a structure is something like the natural numbers $\mathcal{N}$. At the crucial step (A) of the Liar inference, the relation *Exp* becomes salient, and so the salience structure shifts to $\langle \mathcal{N}, \mathit{Exp} \rangle$ at (B). According to the theory I shall develop, *Exp* is significantly more complex than anything in $\mathcal{N}$, so this is a genuine change of context. I shall show how this change can lead to an expansion of the domain of truth conditions for the expanded context, thus solving the expansion problem. This in turn, will lead to a context-based resolution of the paradox.

So far, I have suggested that there is good linguistic reason to take salience structure to be part of context, and to see salience structure as changing as the discourse progresses. I have thus argued that we should see salience structure as playing a role in *ordinary* context dependence. I have also argued salience structure changes in the course of the Liar inference. I shall argue in a moment that together with the semantics of expressions like Exp, it provides the resources for understanding the extraordinary context shifts involved in the Liar.

1.2.2 Truth Conditions and Context

Recall from Section (1.1.4), the expansion problem presented us with two tasks. We have now done the second: explaining how there can be a context shift between steps (A) and (B) in the Liar inference. This leaves the first: explaining how this context shift can induce an expansion of the background domain of truth conditions from which propositions are constructed. Once

[17]Of course, I am idealizing away from the many difficulties involved in applying ideas about ordinary discourse to *proofs*.

this is done, we will have explained the extraordinary context dependence of the propositional quantifier $\exists p$. This in turn will show how the Liar sentence is context dependent, in the right way to resolve the paradox.

Before turning to the remaining task, it will be useful to make one preliminary point about the aspect of context I identified in Section (1.2.1). Salience structures, it should be noted, display much of the behavior of *situations* I mentioned in Section (1.1.5). Salience structures are structured and partial, and provide a representation of a part of the actual world of utterance (or at least, how speakers take the world of utterance to be). Like the situation-theoretic approach had it, salience structures expand in a very simple way in the course of the Liar inference, by adding the relation *Exp*.

In section (1.1.5), I rejected a direct approach to solving the expansion problem based on situations. In light of this, I should stress that salience structures, as I see them, are a representation of a very particular aspect of context. They provide a running list of items salient in a discourse at a given point. As an aspect of *context*, this information helps to determine what truth conditions are expressed by an utterance. But contexts in general, and salience structures in particular, are not truth conditions themselves. Thus, the situation-like behavior we have observed is has not been offered as behavior of the domain of truth conditions. One way to put the problem we now face is to explain why the domain of truth conditions really does inherit some expandability from this situation-like behavior of context.[18]

To this end, let us look at a domain of truth conditions, say W (for worlds). The usual custom in formal semantics is to take W to be a set of indices, assumed to be rich enough to do whatever is needed to accurately model truth conditions. Though it is usually assumed to be infinite, the extent of W is usually left undescribed, as it is usually not an issue. In practice, we usually just use W to specify something like $\{w \in W \mid \varphi(w)\}$, and say no more. Or, we work with toy examples, such as we generate when we say, for instance, 'consider two worlds: in w_1 A happens and in w_2, B happens'. For most purposes, this is fine. But if we are to investigate the extraordinary context dependence of this domain, we cannot leave the matter in this state.

[18] The situation-theoretic approach of Barwise and Etchemendy merges aspects of context and content in a way that makes comparisons difficult. In a situation-theoretic proposition of the form $\{s; [\sigma]\}$, the situation s plays the role of 'what the utterance is about'. This may be similar to the notion of topic which motivates salience structure. But the motivations from topic are not directed towards a constituent of a proposition, but rather the discourse setting in which the utterance appears. Hence, it by no means clear that their notion is the same as mine, or even a notion of context at all.

There are some ideas available about what a domain of worlds W should look like. For instance, recall from the early days of modal logic the idea that a possible world be thought of as a maximal-consistent set of sentences.[19] Now, as a philosophical thesis, this idea has been subjected to some rather harsh criticism. But it does have one nice technical feature. If we wish to use possible worlds to do semantics, the maximal consistent sets of sentences give us precisely the right distribution of worlds. First of all, it gives us enough worlds. For any distinct things speakers can express, there are worlds to mark the difference. Moreover, there are no extra worlds, in that there are no worlds that cannot be distinguished by what speakers can say.

These days, the maximal-consistent set of sentences approach to possible worlds is not very widely accepted. But I believe the point about the distribution of worlds with respect to what speakers can say carries over to any approach to possible worlds that makes them appropriate for comprising the contents of utterances—any approach to truth conditions. This is suggested by some fundamental ideas about assertion and content.

It is a familiar idea that in expressing a proposition in a context, a speaker fundamentally divides possibilities. The content of the speaker's assertion—the proposition expressed—precisely reflects this division. The possibilities are divided into the class of those in which the claim is true, and the class in which it is not true. The proposition expressed is the former class.[20] But asserting a proposition is not by any means a matter of surveying a fixed and independently given domain of options, and indicating which are in the proposition and which are out. Outside of philosophy examples, we almost never do this. When we wish to specify the worlds in the proposition expressed by, say, 'Grass is green', we can do little except say that they are the worlds in which grass is green. Even when we do have some specified list of possibilities to chose from, making arbitrary choices

[19]Carnap [6] makes use of 'state descriptions': maximal-consistent sets of atomic and negated atomic sentences. A standard example of the straightforward maximal-consistent set of sentences view is Jeffrey [25].

[20]This idea is given elegant expression in Stalnaker [51]. Once again, it is worth noting that as the relevant issue here is only one of truth conditions, I am acquiescing to the view that contents of utterances are truth conditions. Nothing about the arguments I present here would be affected by viewing contents as more fine-grained, or by issues about the contents of other speech acts than assertion.

I have also (sometimes) acquiesced to the widely held view that individual truth conditions are possible worlds. Precisely what objects truth conditions are does not matter for my arguments here, nor does it matter if exactly the same notion is used to explain modality and content. What does matter is that truth conditions, whatever they are, must be part of an accurate theory of content.

among them does not constitute making a normal assertion—expressing a proposition through the use of language. We are often hard-pressed to come up with a sentence which expresses the division of possibilities we have made in such an exercise.

In expressing propositions, speakers rely on linguistic resources to effect a division of worlds. But regardless of what we think the right theory of possible worlds, they do so without the benefit of a fixed menu of worlds from which to make choices. Rather, they rely on linguistic resources both to identify worlds and mark them as either in or out of a proposition. This is not really so surprising. But if it is right, then we should not expect the worlds involved in any expressed proposition to go beyond those which can be differentiated by linguistic means. We thus come to expect the same distribution of worlds as we saw with the maximal-consistent set of sentences view. We should not expect to have more worlds involved in propositions than can be distinguished by what speakers can say. And of course, we must have at least as many worlds as can be so distinguished.

We must thus see the domain of truth conditions—worlds—as constrained by the resources speakers have for expressing propositions. These resources may change as context changes. Hence, we must always see the domain of truth conditions available for the construction of contents in a given context as restricted by what speakers can express in that context. In light of what we just observed about assertion, this may have the ring of a truism. I do not think it is such a surprising claim, but it is not a truism; for the claim is not just that what speakers do express is limited by what they can express. Rather, the claim is that the domain of truth conditions available in a context at all, from which any propositions expressed must be constructed, is so limited. Hence, the quantifier $\exists p$ winds up being implicitly restricted by context.[21]

We now have some idea what the extraordinary context dependence involved in the Liar is. It is the dependence of the background domain of truth conditions upon context. This is not ordinary context dependence, as it is the context dependence of the *background* domain, the maximal domain of truth conditions with which speakers may form propositions. It markedly does not behave like other forms of context dependence. But it is a form of context dependence nonetheless.

This shows how there can be extraordinary context dependence at work in the Liar. But to resolve the paradox, we need more. We need to show that the background domain of truth conditions is not only context dependent,

[21] An alternative argument for a view like this can be found in Stalnaker [52].

but that it can *expand* as context shifts. Especially, it must expand as a salience structure expands by the addition of certain semantic relations.

To fully address this, I shall have to turn to some formal apparatus, which will be discussed at length in a Part (2). But first, let me give a brief informal sketch of why this is the result we should expect.

Speakers can describe and individuate very complicated truth conditions. This is especially true in languages which include semantic predicates, like truth or expression. As is well-known, in these languages, speakers can simulate *infinitely long sentences*. Recall, for instance, the oft-noted example of the doctrine of Papal infallibility: 'Everything the pope says is true' amounts to the (potentially) infinitary conjunction $\bigwedge_s \text{PopeSays}(\ulcorner s \urcorner) \rightarrow s$. Hence, speakers can make claims and individuate worlds in ways whose *complexity* mimics such infinitary claims.

It is because of this that extraordinary context dependence really matters. When asking about which such infinitely complex claims speakers can use to individuate worlds, we have to pay attention to the complexity of the infinite sets involved. In the Papal infallibility example, for instance, speakers crucially rely upon the predicate *PopeSays* to specify the collection to be conjoined. We should expect that if context itself provides more complex concepts, say through a salience structure, speakers will be able to use these to formulate more complex infinitary descriptions. If the descriptions get complex enough, they will lead to a genuine expansion of the expressive resources speakers can bring to bear in a context. As these resources determine the extend of the background domain of worlds, such an expansion of expressive resources can amount to a genuine expansion of the domain of truth conditions they can use to express propositions.

Though I do not want to endorse the maximal-consistent set of sentences view of possible worlds, this situation is made vivid by it. Suppose, at least, that the distribution of worlds looks like the maximal-consistent set view has it. When the language involves semantic expressions, I have been arguing, we in effect have to deal with something like infinitary sentences. Which infinitary sentences? Those in which the infinitary operations correspond to resources speakers have available in a context. Which are those? Context will have to tell us. Some contexts will provide resources for far more complex infinitary operations than others. These contexts will allow for significantly more infinitary sentences to go into maximal-consistent sets, and so allow speakers to describe significantly more complicated circumstances. I shall show below in Part (2) that we really do get genuinely new possibilities out of these expanded sets of sentences. Though not all that surprising, this result is not entirely trivial either.

To explore the details of this idea, we will have to turn to some more sophisticated formal apparatus, to be developed in Part (2). But let me first review what this tells us about extraordinary context dependence and the expansion problem. I have argued that we can see the Liar sentence as exhibiting extraordinary context dependence in the domain of the quantifier $\exists p$, which in turn amounts to extraordinary context dependence of the background domain of truth conditions available to speakers to construct propositions in a given context. I isolated a shift in context in the Liar inference: a shift in salience structure. The salience structure shifts by the addition of a semantic relation to it. This, I have argued, leads to an expansion of the background domain of truth conditions. It does so, as this domain is dependent upon speakers' resources for describing or individuating truth conditions. As is illustrated by the maximal-consistent set of sentences idea, if context provides complex enough semantic relations, these resources expand, allowing for an expansion of the domain of truth conditions.

This expansion allows for the coherence of the claims made at the crucial points in the Liar inference. At (A) of Section (1.1.3), it is claimed there is no proposition for the Liar sentence l to express. Relative to the background domain of this context, this claim is true. But this induces an expansion of the salience structure, which in turn induces an expansion of the background domain of truth conditions. The result is there are more propositions available. Thus, there is a proposition for the l to express, making the claim at (B) true.

2 A Formal Model of Expansion

We now have a rough sketch of how the expansion problem may be approached. We have some good linguistic and philosophical motivations for seeing extraordinary context dependence in the Liar sentence, and we have a rough-and-ready sketch of how it works. But we are lacking in details. The details are especially important, as we still have not really seen why the kind of change in expressive resources I discussed in Section (1.2.2) really does lead to an *expansion* of the domain of truth conditions.

In this part, I shall fill in the details by developing a formal model of extraordinary context dependence and expansion. In Section (2.1) I shall briefly provide some background. I shall then introduce the formal machinery for working with salience structures as contexts in Section (2.2). An account of what domain of truth conditions should be attached to a salience structure will be provided in Section (2.3). My goal is then to show that as context changes in the course of the Liar inference, this domain of truth conditions expands. This will be done in Section (2.5). But before it can be accomplished, the sorts of semantic relations involved in the Liar must be given reasonable semantic values. This will be the task of Section (2.4). Section (2.6) will explain more formally the role of infinitary languages I mentioned above in Section (1.2.2). Finally, in Section (2.7) I shall review how the formal model developed solves the expansion problem.

I should stress that what I am about to do is work out formally the details of one relatively simple case. This will not be a fully general theory of context dependence, or even of all aspects of extraordinary context dependence. Such a general theory would be a massive undertaking, and beyond the scope of this paper, if not our current state of knowledge. But in working out a simple case in full detail, I hope to give some more substance to the rather vague sketch I gave above, as well as illustrate how the development of a more comprehensive theory may proceed. I shall offer, as they say, 'proof of concept', even if not the full production model.

2.1 Some Background

In this part of the paper, I shall switch attention from motivations in linguistics to mathematical logic. As the informal discussion of Part (1) shows, we need to explain what speakers can express in a given context. I shall do so by way of the tools of *definability* theory, in particular, the theory of admissible sets with urelements of Barwise [1]. We want some uniform explanation of what set of worlds a given salience structure determines. The answer will be provided by identifying the right admissible set construction over an arbitrary structure, and then applying it to the salience structures context provides. In fact, I shall argue that for a given salience structure $\mathfrak{M}$, the right domain is the one provided by the next admissible $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. This will play the role of fixing 'expressive resources available to speakers' in a context. What speakers can express is limited by what can be done in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. What worlds are available to form propositions corresponds to the structures in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$.

In the course of demonstrating how the domain of worlds expands as $\mathfrak{M}$ expands, I shall need to make use of some fairly fine-grained complexity measures. I shall not rely upon the maximal-consistent set of sentences approach to possible worlds I mentioned in Section (1.2.2). In what follows, worlds will be represented by structures in the usual model-theoretic sense. Even so, infinitary logic will provide some of the tools for making complexity measurements and describing structures. Technically, other tools of definability theory can do this job just as well. But there are two reasons for using infinitary logic. One purely technical reason is that infinitary logic makes for some reasonably elegant results, and allows me to borrow some important results from model theory without modification. On the philosophical side, however, I have already noted in Section (1.2.2) the rough match-up between languages containing semantic expressions and infinitary languages. This rough match-up will help to make more vivid how to understand the technical apparatus I shall develop below, and will provide a guide for identifying domains of truth conditions. In Theorem (2.6.1) and Remark (3.3.12), I shall show that this match-up is more than a rough intuitive guide: it can be given precise content.

In the following, I shall assume familiarity with the fundamental works of Barwise [1] and Moschovakis [40]. I shall try as much as possible to follow the notation of Barwise. A couple of specific points are worth mentioning. The universe of a structure $\mathfrak{M}$ is written M. Admissible sets are denoted by blackboard Bold characters. $\mathbb{A}_{\mathfrak{M}}$ is an admissible over $\mathfrak{M}$, i.e. a structure $\langle \mathfrak{M}, A, \in, \ldots \rangle$ with $M \cup A$ transitive which models KPU. The most important admissibles we shall encounter below are of the form $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$: the smallest admissible $\mathbb{A}_{\mathfrak{M}}$ containing M. As blackboard bold letters are taken, $\mathcal{N}$ denotes the natural numbers. I shall continue the practice of writing syntactic items in $\mathsf{sans\ serif}$, and their interpretations in *italics*. When introducing a term which defines a relation already given, I shall follow the usual custom of putting a dot over it. Gödel codes for formulas are written $\ulcorner \varphi \urcorner$, though I shall continue the practice of suppressing quotation marks where possible. I should also mention the terminological conventions of generalized recursion theory that a relation is $\mathbb{A}$-recursive if it is Δ_1 on $\mathbb{A}$, $\mathbb{A}$-r.e. if it is Σ_1 on $\mathbb{A}$, and $\mathbb{A}$-finite if it is in $\mathbb{A}$. A function is $\mathbb{A}$-recursive if it has an $\mathbb{A}$-r.e. graph.

2.2 Context and $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$

A context, I argued in Section (1.2.1), provides a salience structure. I have also argued that as our concern is a specific sort of extraordinary context dependence, we may ignore other aspects of context. Hence, from now on, I shall take a context simply to be a salience structure.

I have already noted that we should assume salience structures—contexts—contain basic syntactic resources. They should thus resemble closely enough the natural numbers. But, we cannot restrict our attention to $\mathcal{N}$, as we need to allow contexts to change. We need to be able to consider many distinct structures, each of which is 'close enough' to $\mathcal{N}$.

The notion of 'close enough to the natural numbers' is given by

Moschovakis's notion of an *acceptable structure* [40, 5A]. $\mathfrak{M}$ is acceptable if it contains *elementary* operations for coding finite sequences from M, including an elementary ordering $\mathcal{N}^{\mathfrak{M}}$ isomorphic to the natural numbers. Recall that for acceptable $\mathfrak{M}$, the relations on M in $\mathbb{H}\text{YP}_{\mathfrak{M}}$ are exactly the hyperelementary relations. For the case of the natural numbers $\mathcal{N}$, these are the hyperarithmetic relations.

Let $\mathcal{L}$ be the language of an acceptable $\mathfrak{M}$. Assuming $\mathcal{L}$ has a countable signature, it can be coded up in $\mathcal{N}^{\mathfrak{M}}$. It will be useful to make the stronger assumption that $\mathcal{K} = \mathcal{L} \cup \{\dot{\mathsf{m}} \mid m \in M\}$ can be coded in $\mathcal{N}^{\mathfrak{M}}$. This amounts to assuming that M can be coded in $\mathcal{N}^{\mathfrak{M}}$. We need not require this coding to be elementary—hyperelementary suffices.

Definition 2.2.1. A structure $\mathfrak{M}$ is *strongly acceptable* if it is acceptable and there is a hyperelementary coding of M in $\mathcal{N}^{\mathfrak{M}}$.

Observe that strong acceptability implies *countability.*

From now on, I shall assume a context is a strongly acceptable structure, with finitely many relations. For our intended application to the Liar, we need only structures like $\mathcal{N}$, $\langle \mathcal{N}, Exp \rangle, \ldots$, all of which meet these conditions. (For most of what follows, we could have allowed the coding apparatus to be only hyperelementary: what Moschovakis calls 'almost acceptability'. As as our intended examples are acceptable, however, we may use the stronger condition.) Where convenient, I shall also assume that contexts do not contain functions, though this will not really matter much.

The utility of strong acceptability flows from the following lemma.

Lemma 2.2.2. *Suppose $\mathfrak{M}$ is strongly acceptable.*

1. $\mathbb{H}\text{YP}_{\mathfrak{M}}$ *is projectible into $\mathcal{N}^{\mathfrak{M}}$ by a univalent notation system.*

2. *(Selection.) If $R(x,y)$ is a $\mathbb{H}\text{YP}_{\mathfrak{M}}$-recursive relation, there is a $\mathbb{H}\text{YP}_{\mathfrak{M}}$-recursive function*
$$f(x) = \mu_\pi y R(x,y)$$
which returns the y with least code in $\langle \mathcal{N}^{\mathfrak{M}}, \leq^{\mathfrak{M}} \rangle$.

Sketch of proof. For (1), the standard proof of the projectibility of $\mathbb{H}\text{YP}_{\mathfrak{M}}$ into $\mathfrak{M}$ given in [1, VI.4.12] nearly suffices. We just observe that strong acceptability allows us to choose codes in $\mathcal{N}^{\mathfrak{M}}$. Let π be the notation system so given. Then we may define $\pi'(x) =$ the least y such that $y \in \pi(x)$. As the ordering of the coding structure is elementary, π' is still $\mathbb{H}\text{YP}_{\mathfrak{M}}$-recursive.

(2) follows immediately from (1). ■

It follows from Lemma (2.2.2) that for a strongly acceptable $\mathfrak{M}$, $\mathbb{HYP}_{\mathfrak{M}}$ has most of the nice properties of $\mathbb{HYP}_{\mathcal{N}}$. It is countable, locally countable, resolvable, recursively listed, and projectible into $\mathcal{N}^{\mathfrak{M}}$.

2.3 Truth Conditions in $\mathbb{HYP}_{\mathfrak{M}}$

Now, that we have contexts in hand, I need to show how to construct an appropriate domain of truth conditions relative to a context. I shall show how given a context $\mathfrak{M}$, $\mathbb{HYP}_{\mathfrak{M}}$ provides a good model of the expressive resources speakers can bring to bear in that context. In particular, it provides the domain $W_{\mathfrak{M}}$ of truth conditions or worlds available in $\mathfrak{M}$.

Following the usual practice, I shall identify truth conditions with *structures.* We could of course use any number of devices instead (Hintikka sets, for instance), but this convention allows us to make use of some results from model theory.

We do need to be a bit careful about languages. Let us fix an initial context: a strongly acceptable structure $\mathfrak{M}$ with language $\mathcal{L}$. Let us assume that $\mathcal{L}$ contains no semantic vocabulary. In a moment, we will augment $\mathcal{L}$ to an $\mathcal{L}^+$ which contains appropriate semantic terms. $\mathcal{L}$ is the language of the context $\mathfrak{M}$, and we should generally expect this to be a proper subset of the language available to speakers. $\mathcal{L}$'s signature is the *salient* expressions only. However, in what follows, the interesting distinction will be between whatever language speakers may employ without semantic terms, and that with semantic terms. With this in mind, it will be a useful simplification to identify the language of the initial context with the non-semantic component of the language available to speakers. Thus, the relevant 'worlds' are $\mathcal{L}$-structures.

In Section (1.2.2), I proposed we think of the truth conditions available in a context as corresponding to what speakers can describe or individuate. As I mentioned there, there is a rough match-up between the expressive capacities of speakers of languages with semantic relations and infinitary languages, and that it is up to context to fix the strength of the appropriate infinitary language. As I mentioned, this will be made more precise in Theorem (2.6.1) and Remark (3.3.12), but before this can be done, we need to investigate these languages and their models in more detail. The infinitary languages we need will be fragments of the infinitary language $\mathcal{L}_{\infty\omega}$, which allows arbitrary conjunctions and disjunctions, but only finite quantifier prefixes. Recall, a fragment is basically a subset of $\mathcal{L}_{\infty\omega}$ which contains all finitary formulas, and is closed under substitutions of terms, subformulas, and finitary operators. Nice fragments of $\mathcal{L}_{\infty\omega}$ are *admissible fragments.* An admissible fragment $\mathcal{L}_{\mathbb{A}} = \{\varphi \in \mathcal{L}_{\infty\omega} \mid \varphi \in \mathbb{A}\}$. (See [1] or [28] for details.)

The basic idea is to use $\mathbb{HYP}_{\mathfrak{M}}$ to capture the expressive resources available in $\mathfrak{M}$. Thus, when looking for infinitary languages speakers can use to describe worlds, we should ask about fragments $\mathcal{L}_B$ that are in $\mathbb{HYP}_{\mathfrak{M}}$. Worlds should then correspond to complete consistent theories that can be formulated in these fragments. We need to show that structures in $\mathbb{HYP}_{\mathfrak{M}}$ correspond appropriately to these theories.

The essential result which allows this project to go forward is due to Nadel, and it tells us that in fact any such theory has a model in $\mathbb{HYP}_{\mathfrak{M}}$.

Theorem 2.3.1 (Nadel [41]). *Let $\mathcal{L}_B$ be a countable fragment of $\mathcal{L}_{\infty\omega}$. Suppose $\mathbb{A}$ is admissible and $\mathcal{L}_B \in \mathbb{A}$ is countable in the sense of $\mathbb{A}$. Let $T \in \mathbb{A}$ be a* complete *consistent $\mathcal{L}_B$-theory. Then T has a model in $\mathbb{A}$.*

Remark 2.3.2.

1. The local countability of $\mathbb{HYP}_{\mathfrak{M}}$ for strongly acceptable $\mathfrak{M}$ means we may consider any $\mathcal{L}_B \in \mathbb{HYP}_{\mathfrak{M}}$.

2. The assumption of completeness cannot be removed, even for the special case of $\mathbb{HYP}_{\mathcal{N}}$. (See the discussion of Nadel [42].)

3. The proof of Theorem (2.3.1) proceeds by showing that a reasonably careful version of a canonical model construction can be carried out in $\mathbb{A}$. In fact, some more complex constructions, such as the proof of the $\mathfrak{M}$-completeness theorem or the forcing proof of the omitting types theorem of [29] can be carried out within $\mathbb{A}$. Thus, if we thought that 'genuine worlds' should have a standard part, we could still find as many as we may need in $\mathbb{HYP}_{\mathfrak{M}}$.

Now, this gives us good reason to think that $\mathbb{HYP}_{\mathfrak{M}}$ contains enough structures to reflect what speakers can express in context $\mathfrak{M}$. In Section (1.2.2) I also suggested that an appropriate domain of truth conditions for a given context should not contain too many truth conditions. Any difference in worlds in the domain should correspond to something speakers can express. This is indeed what we find, as is shown by another result of Nadel.

Theorem 2.3.3 (Nadel Basis Theorem [42]). *Let $\mathfrak{B}$ and $\mathfrak{C}$ be structures in an admissible $\mathbb{A}$. Then $\mathfrak{B} \equiv \mathfrak{C}(\mathcal{L}_{\mathbb{A}})$ implies $\mathfrak{B} \equiv_{\infty\omega} \mathfrak{C}$.*

By a well-known theorem of Karp (see, e.g. [1]), $\mathfrak{B} \equiv_{\infty\omega} \mathfrak{C}$ iff $\mathfrak{B}$ and $\mathfrak{C}$ are partially isomorphic (have the back-and-forth property). For countable structures, partial isomorphism implies isomorphism, so we have

Corollary 2.3.4. *Let* countable *structures* $\mathfrak{B}$ *and* $\mathfrak{C}$ *be elements of an admissible* $\mathbb{A}$. *Then* $\mathfrak{B} \equiv \mathfrak{C}(\mathcal{L}_{\mathbb{A}})$ *implies* $\mathfrak{B} \cong \mathfrak{C}$.

The countability of $\mathbb{HYP}_{\mathfrak{M}}$ for strongly acceptable $\mathfrak{M}$ ensures that for $\mathfrak{B}, \mathfrak{C} \in \mathbb{HYP}_{\mathfrak{M}}$, $\mathfrak{B} \equiv \mathfrak{C}(\mathcal{L}_{\mathbb{HYP}_{\mathfrak{M}}})$ implies $\mathfrak{B} \cong \mathfrak{C}$.

It follows that for any distinct structures in $\mathbb{HYP}_{\mathfrak{M}}$, there is something that speakers can express in $\mathfrak{M}$ that marks the difference. Any two non-isomorphic structures do not make all the same $\mathcal{L}_{\mathbb{HYP}_{\mathfrak{M}}}$-sentences true. For any sentence φ on which they differ, we can construct a fragment $\mathcal{L}_A$ in $\mathbb{HYP}_{\mathfrak{M}}$ which contains φ. As $\models$ is Δ in KPU, the $\mathcal{L}_A$-theories of the structures are in $\mathbb{HYP}_{\mathfrak{M}}$. These $\mathcal{L}_A$-theories differ, as one contains φ and the other $\neg\varphi$. Hence they provide something in $\mathbb{HYP}_{\mathfrak{M}}$, something speakers in the context can express, which witnesses the difference between the two structures.

$\mathbb{HYP}_{\mathfrak{M}}$ thus recommends itself as the domain from which to draw truth conditions. There are enough structures in $\mathbb{HYP}_{\mathfrak{M}}$ to model what speakers can express in context $\mathfrak{M}$, and not too many. It provides just the right truth conditions. We might wonder if we really do need all of $\mathbb{HYP}_{\mathfrak{M}}$ to do the job, or if some piece of it might suffice. We do need all of it, as I shall show in Section (2.5). As I mentioned, the question of how closely languages with semantic expressions match up to the infinitary fragments I have used in discussing $\mathbb{HYP}_{\mathfrak{M}}$ will be addressed in Theorem (2.6.1) and Remark (3.3.12).

Let me end this section with one useful technical point. Above, we looked at structures in $\mathbb{HYP}_{\mathfrak{M}}$. The strong acceptability of $\mathfrak{M}$ allows us to assume that these structures in fact live on M, the domain of $\mathfrak{M}$. This fact will make it more convenient to invoke the machinery of inductive definitions on $\mathfrak{M}$ in what follows.

Lemma 2.3.5. *Let* $\mathfrak{A} = \langle A, R_1, \ldots, R_n \rangle$ *be a structure in* $\mathbb{HYP}_{\mathfrak{M}}$. *Then there is a structure* $\mathfrak{B} \in \mathbb{HYP}_{\mathfrak{M}}$ *such that* $B \subseteq M$ *and* $\mathfrak{A} \cong \mathfrak{B}$. *Moreover, there is an isomorphism* $f \colon \mathfrak{A} \cong \mathfrak{B}$ *in* $\mathbb{HYP}_{\mathfrak{M}}$.

Proof. The structure $\mathfrak{B}$ is essentially the structure generated by the codes of elements of $\mathfrak{A}$. Let π be a univalent notation system for $\mathbb{HYP}_{\mathfrak{M}}$. Then let $B = \{\pi(a) \mid a \in A\}$, $S_i = \{\langle \pi(x), \pi(y) \rangle \mid R_i(x, y)\}$. Clearly π induces an isomorphism $\bar{\pi} \colon \mathfrak{A} \cong \mathfrak{B}$.

It remains to be shown that each of B, S_i and $\bar{\pi}$ is in $\mathbb{HYP}_{\mathfrak{M}}$. Observe that each A and R_i is in $\mathbb{HYP}_{\mathfrak{M}}$ as $\mathfrak{A}$ is. As π is a total recursive function, it has a Δ graph. Hence, B is Δ in $\mathbb{HYP}_{\mathfrak{M}}$ on M, and each S_i is Δ on the relevant M^k. Thus, by Δ-separation, each component of $\mathfrak{B}$ is in $\mathbb{HYP}_{\mathfrak{M}}$, so by closure under pairing, $\mathfrak{B}$ is. Finally, as $A \times B$ is in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$, so is $\bar{\pi}$, again by Δ-separation. ■

Finally, recall that in looking at the dynamics of context, we will look at sequences of salience structures like $\mathfrak{M}, \langle \mathfrak{M}, \mathit{Exp} \rangle, \ldots$. It will be useful at some points to also consider structures like $\langle \mathfrak{B}, \mathit{Exp} \rangle$ for a given $\mathfrak{B} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. This only makes sense if $\mathfrak{B}$ includes in its domain the relevant codes for sentences and worlds. Remark (2.3.2) shows this can be arranged. However, to simplify matters somewhat, I shall assume where necessary that we are talking only about structures with domain M. $|\mathfrak{B}| = M$ is hyperelementary, so this does not change any complexity issues. Assuming this is essentially to make the assumption one often sees in discussions of quantified modal logic, that there are no contingent domains. This assumption is in some ways inaccurate, but none of the reasons it is are relevant here. Hence, without comment, I shall assume it where necessary.

2.4 A Parameterized Kripke Construction

We now have at our disposal a context, given by a strongly acceptable $\mathfrak{M}$. $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ models the expressive resources of $\mathfrak{M}$, and provides a domain of truth conditions for the context, given by the structures in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. This gives us a basic formal model of the extraordinary context dependence of the domain of truth conditions upon context.

We need to show that as the context changes by adding a semantic relation like *Exp* to the salience structure, we genuinely expand the domain of truth conditions. This requires showing that the universe $\mathbb{H}\mathrm{YP}_{\langle \mathfrak{M}, \mathit{Exp} \rangle}$ is indeed larger than $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$, and most importantly, that it is larger in such a way as to include genuinely more models—more truth conditions. This will essentially solve the expansion problem.

To do this, we need to say something about the *semantics* of the appropriate semantic relations. Before turning to this, we should look once more at what semantic relations are relevant.

In the discussion of Part (1), I used a unary truth predicate of propositions $\mathit{Tr}(p)$ and the expression relation *Exp*. We observed there that because ordinary context dependence is not at issue, we could take *Exp* to be the two-place $\mathit{Exp}(s, p)$. We also saw that, for the extraordinary context dependence of the Liar, the relevant aspect of context is given by a salience structure $\mathfrak{M}$, and the only relevant effect of context is to fix a domain of truth conditions $W_{\mathfrak{M}}$, which we now have identified as the domain of structures in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$.

Under these assumptions, we can make a further simplification in our semantic relations. Consider the binary relation $Tr(w, s)$ which holds iff sentence s is true in world w. Given our assumptions, once a domain $W_{\mathfrak{M}}$ is fixed, this relation suffices to fix *Exp* as well.

Remark 2.4.1. For context $\mathfrak{M}$ with domain of truth conditions $W_{\mathfrak{M}}$, $Exp(s, p) \leftrightarrow p = \{w \in W_{\mathfrak{M}} \mid Tr(w, s)\}$.

(I shall return to the issue of how to recover *Exp* from *Tr* in Section (3.4).) In light of this, in what follows, I shall work exclusively with the binary *Tr*. $Tr(w, s)$ may be thought of as a *parameterized* version of *Exp*. Technically, it is convenient to make use of the parameterized version, as we have no hope of coding up all sets of structures on a countable $\mathfrak{M}$.

Now, let us fix some some syntax to go along with this observation. As we have been doing, fix a strongly acceptable $\mathfrak{M}$ with language $\mathcal{L}$. Again let $\mathcal{K} = \mathcal{L} \cup \{\dot{\mathsf{m}} \mid m \in M\}$. (We may treat $\mathfrak{M}$ as a $\mathcal{K}$-structure as well.) The assumption that every element of M has a name smoothes over some technicalities, so I shall restrict attention to $\mathcal{K}$, and note the assumption that structures in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ are taken to have domain M. As I noted in Section (2.3), what is of importance here is not so much the language with which we start as what happens when semantic relations are added to it. Hence, we will be concerned with the language $\mathcal{K}^+$ that extends $\mathcal{K}$ by adding a binary relation symbol Tr. On its intended interpretation, $\mathsf{Tr}(x, y)$ holds if y is a sentence of $\mathcal{K}^+$ and x is a a structure in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ and y is true in x. Our task is to build an appropriate *Tr* to interpret Tr for context $\mathfrak{M}$.

A couple of coding devices will be necessary. I shall make use of the elementary function $\dot{\dot{x}}$ which returns the term $\dot{\mathsf{m}}$ for m. Conversely, I shall make use of the elementary function *Den*, which takes a closed term of $\mathcal{K}$ and returns the object it denotes. $Den(\dot{\mathsf{m}}) = m$. As *Den* is elementary, there is a $\mathcal{K}$-expression $\dot{\mathsf{Den}}$ which represents it. (For the proof that *Den* is elementary, and a nice survey of Tarski-type results, see McGee [38].)

Aside from syntax coding, we will need notations for elements of $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ in $\mathfrak{M}$. Lemma (2.2.2) guarantees that for a strongly acceptable $\mathfrak{M}$, $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ has a univalent notation system. This will always be denoted π. The set of notations is D_π, and notation y codes $|y|_\pi$. Using the notation system, we can code structures in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. It is easy to see that the predicate 'is a structure' is inductive on $\mathfrak{M}$. 'Is a structure' is Δ in KPU. Set $Str(x) \leftrightarrow (D_\pi(x) \wedge |x|_\pi$ is a structure). D_π is Σ_1 on $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$, so *Str* is inductive on $\mathfrak{M}$.

Using *Str*, we can carry out a version of the familiar construction of Kripke [32] for Tr. In Kripke's original presentation, and in many others, a partial predicate is constructed; that is, a pair $\langle E, A\rangle$ consisting of an extension and an anti-extension. In many cases this is more information than we need. If we are after the minimal fixed point of the original Kripke construction, for example, the anti-extension A is just $\{\ulcorner\varphi\urcorner \mid \ulcorner\neg\varphi\urcorner \in E\} \cup \{\text{nonsentences}\}$. As we shall not make use of unusual fixed points, we can avoid the added nuisance of keeping track of both extensions and anti-extensions altogether, so long as the difference between the negation of a sentence being in an extension and the sentence not being in the extension is respected. To help do this, we may use a trick due to Gilmore [13] and Feferman [9]. For a formula φ in the language $\mathcal{K}^+$, write φ^* for the result of replacing each negative occurrence of $\mathsf{Tr}(x, y)$ by $\dot{\mathsf{Sent}}(y) \wedge \neg\mathsf{Tr}(x, \dot{\neg}(y))$. Tr always occurs positively in φ^*. Hence, as long as we evaluate φ^* for truth in a structure rather than φ, we will respect the difference we need to.

What should the extension of Tr be? Following Kripke, we can build up a relation in stages.

Definition 2.4.2. Let:

$$Tr_0 = \{\langle x, y\rangle \mid Str(x) \wedge Sent(y) \wedge \langle x, \emptyset\rangle \models y^*\}$$
$$Tr_{\alpha+1} = \{\langle x, y\rangle \mid Str(x) \wedge Sent(y) \wedge \langle x, Tr_\alpha\rangle \models y^*\}$$

(I have suppressed some coding details to keep these readable.) To see that the $*$ trick works, note that $\langle x, Tr_\alpha\rangle \models (\neg\mathsf{Tr}(z, y))^*$ just in case either $\neg Sent(y)$ or $\langle z, \dot{\neg}(y)\rangle \in Tr_\alpha$. Hence, we have the same effect as a partial interpretation with its anti-extension consisting of non-sentences and sentences whose negations are in the extension, evaluated according to the strong Kleene rules.

We wind up getting all of *Str* at the first stage, as any valid sentence of $\mathcal{K}$ is true in each structure. We also get at the first stage all the sentences we ever get from $\mathcal{K}$.

This definition provides us with a monotone operator, which has a least fixed point. Let us call this Tr. One of the important consequences of Tr being a fixed point is given by the following.

Lemma 2.4.3 (The Weak T Property). *For any $\mathcal{K}^+$-sentence φ:*

$$(Tr(x, \ulcorner\varphi\urcorner) \vee Tr(x, \ulcorner\neg\varphi\urcorner)) \rightarrow (Tr(x, \ulcorner\varphi\urcorner) \leftrightarrow \langle x, Tr\rangle \models \varphi).$$

This property makes Tr a plausible interpretation of Tr. As with Kripke-style fixed points generally, it makes Tr come reasonably close to expressing the corresponding genuine semantic notion, $\models$. In particular, it does so allowing a fair bit of self-application of the predicate Tr. It is, I suggest, a plausible starting place for interpreting this expression.

Even so, I should pause to comment on my use of the minimal fixed point to interpret Tr. As it is a fixed point, it satisfies the Weak T Property, and so has a reasonable claim to model semantic phenomena with some degree of accuracy. Of course, there are lots of other fixed points, and many have wondered what reasons there might be to choose one over another (cf. Gupta [22]). I do not intend to take a stand on this issue here. The minimal fixed point is good enough to pursue the formal development of my approach, and that is all I really require. Indeed, the contextual approach minimizes the need for a position on this matter. Semantic expressions can have many different extensions in different contexts, so there is no need to declare in advance of a thorough investigation of various contexts that one extension is *the* right one. If in some particular sort of context, there was good reason to take some other fixed point, that would be fine. For the context shift in the Liar inference, where there is little additional contextual information involved, the minimal fixed point provides an adequate model. That is all we need.

Now, we have an interpretation of the semantic expression Tr of $\mathcal{K}^+$. At the crucial point in the Liar inference, the semantic relation Tr is added to an initial salience structure $\mathfrak{M}$ (as I shall discuss further in Section (3.4)). I claim that this will induce an expanded domain of structures $W_{\langle \mathfrak{M}, Tr\rangle}$ in $\mathbb{H}\mathrm{YP}_{\langle \mathfrak{M}, Tr\rangle}$. To show this, we need to know a little more about the complexity of Tr.

It is a standard result for the usual Kripke construction that its least fixed point is inductive non-hyperelementary (as in Burgess [4] or McGee [38]). The same holds for Tr, but the situation is somewhat more complicated. It is clear from Definition (2.4.2) that Tr is inductive, but the presence of Str requires a little more effort to show that it is inductive on $\mathfrak{M}$.

Lemma 2.4.4. *Tr is inductive on $\mathfrak{M}$.*

Proof. It is well-known that satisfaction is hyperelementary on $\mathfrak{M}$, and in particular that $\mathfrak{A} \models \varphi$ is Δ in KPU in parameter $\mathfrak{A}$. Thus, there is an inductive relation $\dot{\models}$ on $\mathfrak{M}$ such that:

$$c \dot{\models} d \quad \text{iff} \quad Str(c) \wedge Sent_{\mathcal{K}}(d) \wedge \exists \varphi \in \mathcal{K}(d = \ulcorner \varphi \urcorner \wedge |c| \models \varphi).$$

We also know that Str is inductive on $\mathfrak{M}$.

We now need to write out a $(\mathsf{Tr}, \dot{\mathsf{Str}}, \dot{\models})$-positive definition of $Tr(x, y)$, which can be done as follows (taking a hint from McGee [38]):

$$
\begin{aligned}
&\dot{\mathsf{Str}}(x) \wedge \dot{\mathsf{Sent}}(y) \wedge \\
&\qquad \big(\dot{\mathsf{Sent}}_{\mathcal{K}}(y) \wedge x \dot{\models} y \\
&\qquad \vee \exists t \exists u[\dot{\mathsf{Term}}(t) \wedge \dot{\mathsf{Term}}(u) \wedge y \doteq \ulcorner \mathsf{Tr}(t,u) \urcorner \wedge \mathsf{Tr}(\dot{\mathsf{Den}}(t), \dot{\mathsf{Den}}(u))] \\
&\qquad \vee \exists t \exists u[\dot{\mathsf{Term}}(t) \wedge \dot{\mathsf{Term}}(u) \wedge y \doteq \ulcorner \neg \mathsf{Tr}(t,u) \urcorner \wedge \mathsf{Tr}(\dot{\mathsf{Den}}(t), \dot{\neg} \dot{\mathsf{Den}}(u))] \\
&\qquad \vee \exists v \exists w[y \doteq (v \dot{\vee} w) \wedge (\mathsf{Tr}(x,v) \vee \mathsf{Tr}(x,w))] \\
&\qquad \vee \exists v \exists w[y \doteq \dot{\neg}(v \dot{\vee} w) \wedge \mathsf{Tr}(x, \dot{\neg} v) \wedge \mathsf{Tr}(x, \dot{\neg} w)] \\
&\qquad \vee \exists v[y \doteq \dot{\neg}\dot{\neg} v \wedge \mathsf{Tr}(x,v)] \\
&\qquad \vee \exists v \exists w[\dot{\mathsf{Var}}(v) \wedge y \doteq \dot{\exists} v(w) \wedge \exists z(\mathsf{Tr}(x, w(v/\dot{z}))))] \\
&\qquad \vee \exists v \exists w[\dot{\mathsf{Var}}(v) \wedge y \doteq \dot{\neg}\dot{\exists} v(w) \wedge \forall z(\mathsf{Tr}(x, \dot{\neg} w(v/\dot{z}))))]\big)
\end{aligned}
$$

This shows Tr to be inductive on $\langle \mathfrak{M}, Str, \models \rangle$. $\dot{\mathsf{Str}}$ and $\dot{\models}$ appear only positively in the above formula, and are inductive themselves, so we can appeal to the Moschovakis Piggyback Lemma [40, 1.C.1] to conclude that Tr is in fact inductive on $\mathfrak{M}$. ■

It is clear that Tr cannot be elementary, but we have not yet seen that it cannot be hyperelementary. The normal Kripke truth predicate is shown to be non-hyperelementary by showing it to be complete for hyperelementary relations. The problem here is that Tr does not directly provide information about $\mathfrak{M}$. It may be helpful to think about the intended interpretations of these items. $\mathfrak{M}$ is the context. Tr provides information about the domain of truth conditions, which is fixed by the context. But we need to do a little work to extract information about the context from Tr.

What we need to do is say enough about what the context $\mathfrak{M}$ is like to use what Tr tells us about circumstances. This is easy. There is a $\mathcal{K}$-formula $\delta(x)$ which holds just of the atomic or negated atomic sentences of $\mathcal{K}$ true in $\mathfrak{M}$. We give δ basically by listing the relations of $\mathcal{K}$, as Tarski and commentators after him have much discussed. Suppose they are $\mathsf{R}_0, \ldots, \mathsf{R}_n$

(here invoking my assumption that there are no functions). Then set:

$$\begin{aligned}\delta(x) \quad \leftrightarrow \quad & \dot{\mathsf{Sent}}_{\mathcal{K}}(x) \wedge \\ & \big(\exists y_1 \ldots \exists y_j (\dot{\mathsf{Term}}(y_1) \wedge \ldots \wedge \dot{\mathsf{Term}}(y_j) \\ & \wedge x \doteq \ulcorner \mathsf{R}_0(y_1, \ldots, y_j) \urcorner \wedge \mathsf{R}_0(\dot{\mathsf{Den}}(y_1), \ldots, \dot{\mathsf{Den}}(y_j))) \\ & \vee \\ & \vdots \\ & \vee \\ & \exists y_1 \ldots \exists y_k (\dot{\mathsf{Term}}(y_1) \wedge \ldots \wedge \dot{\mathsf{Term}}(y_k) \\ & \wedge x \doteq \ulcorner \mathsf{R}_n(y_1, \ldots, y_k) \urcorner \wedge \mathsf{R}_n(\dot{\mathsf{Den}}(y_1), \ldots, \dot{\mathsf{Den}}(y_k))) \\ & \vee \\ & \exists y_1 \ldots \exists y_j (\dot{\mathsf{Term}}(y_1) \wedge \ldots \wedge \dot{\mathsf{Term}}(y_j) \\ & \wedge x \doteq \ulcorner \neg \mathsf{R}_0(y_1, \ldots, y_j) \urcorner \wedge \neg \mathsf{R}_0(\dot{\mathsf{Den}}(y_1), \ldots, \dot{\mathsf{Den}}(y_j))) \\ & \vee \\ & \vdots \\ & \vee \\ & \exists y_1 \ldots \exists y_k (\dot{\mathsf{Term}}(y_1) \wedge \ldots \wedge \dot{\mathsf{Term}}(y_k) \\ & \wedge x \doteq \ulcorner \neg \mathsf{R}_n(y_1, \ldots, y_k) \urcorner \wedge \neg \mathsf{R}_n(\dot{\mathsf{Den}}(y_1), \ldots, \dot{\mathsf{Den}}(y_k)))\big)\end{aligned}$$

δ holds in $\mathfrak{M}$ just of the members of the diagram of $\mathfrak{M}$. A glance at the proof of Lemma (2.4.4) reveals that what is true in a structure according to *Tr* is determined entirely by its diagram. Hence we have the following lemma.

Lemma 2.4.5. *If* $Diag(\mathfrak{A}) = Diag(\mathfrak{B})$ *then:*

$$\forall s(\mathit{Tr}(\pi(\mathfrak{A}), s) \leftrightarrow \mathit{Tr}(\pi(\mathfrak{B}), s)).$$

Indeed, so long we we assume $\mathfrak{A}$ and $\mathfrak{B}$ have the same universe M, then construed as $\mathcal{K}$-structures $\mathfrak{A}_M$ and $\mathfrak{B}_M$, the antecedent is enough to ensure they are isomorphic. As $\mathfrak{M}$ itself is in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$, we can express the Kripke truth predicate over $\mathfrak{M}$ by using our parameterized truth predicate. Write:

$$\mathsf{TM}(x) \leftrightarrow \forall y[\forall z(\delta(z) \rightarrow \mathsf{Tr}(y, z)) \rightarrow \mathsf{Tr}(y, x)].$$

We can then write TM for the extension of TM in $\langle\mathfrak{M}, Tr\rangle$. Lemma (2.4.5) guarantees that TM is essentially the traditional Kripke truth predicate for $\mathfrak{M}$. It differs in having sentences containing Tr as well as those containing TM, but so long as it has the latter, passing from Tr to TM allows us to recover all the classical results about Kripke truth predicates. In particular, we have the result essentially stated by Kripke [32].

Theorem 2.4.6 (Kripke). *TM is complete for inductive sets on $\mathfrak{M}$, and hence is not coinductive on $\mathfrak{M}$.*

This has the immediate corollary we need.

Corollary 2.4.7. *Tr is not coinductive, and hence not hyperelementary.*

2.5 Expansions

So far, we have identified, for context $\mathfrak{M}$, a universe of truth conditions in $\mathbb{HYP}_{\mathfrak{M}}$. We have also identified the relevant semantic relation Tr, and provided a reasonable semantics for it relative to this universe. We also gained some insight into the complexity of Tr. This will now allow us to look at what happens to the domain of truth conditions (structures) as context (salience structure) shifts from $\mathfrak{M}$ to $\langle\mathfrak{M}, Tr\rangle$. We need to show that this expands $\mathbb{HYP}_{\mathfrak{M}}$ in a way that genuinely adds more structures. This will resolve the expansion problem, which in turn will substantiate the claim that there is extraordinary context dependence in the Liar sentence.

Lemma (2.4.4) and Corollary (2.4.7) show Tr to be inductive non-hyperelementary. The basis for expansion is the well-known effect on $\mathbb{HYP}_{\mathfrak{M}}$ of adding such a relation. The result, due to Moschovakis [40], is a generalization of a result of Spector.

Theorem 2.5.1 (Spector-Moschovakis). *Let $\mathfrak{M}$ be acceptable.*

1. *If P is hyperelementary on $\mathfrak{M}$ then $\mathfrak{M}$ and $\langle\mathfrak{M}, P\rangle$ have the same inductive and hyperelementary relations, and $\mathbb{HYP}_{\mathfrak{M}}$ and $\mathbb{HYP}_{\langle\mathfrak{M},P\rangle}$ have the same universe of sets.*

2. *If P and S are inductive non-hyperelementary, then $\langle\mathfrak{M}, P\rangle$ and $\langle\mathfrak{M}, S\rangle$ have the same inductive and hyperelementary relations, and $\mathbb{HYP}_{\langle\mathfrak{M},P\rangle}$ and $\mathbb{HYP}_{\langle\mathfrak{M},S\rangle}$ have the same universe of sets.*

This shows that adding Tr to the context induces a genuine expansion of $\mathbb{HYP}_{\mathfrak{M}}$. It still needs to be verified that this in fact produces more truth conditions—more structures—in $\mathbb{HYP}_{\langle\mathfrak{M},Tr\rangle}$. We can observe immediately that for any ordinal α in $\mathbb{HYP}_{\langle \mathfrak{M}, Tr \rangle}$ but not in $\mathbb{HYP}_{\mathfrak{M}}$, there is an ordering $\langle M, \leq \rangle$ in $\mathbb{HYP}_{\langle \mathfrak{M}, Tr \rangle}$ which is not order-isomorphic to any ordering in $\mathbb{HYP}_{\mathfrak{M}}$. To do full justice to the promise made in Section (1.2.2), however, we need to do more. We need to make clear how the resources of $\mathbb{HYP}_{\langle \mathfrak{M}, Tr \rangle}$ allow speakers to express more in the expanded context. In particular, we need to find a natural way to build structures in $\mathbb{HYP}_{\langle \mathfrak{M}, Tr \rangle}$ whose theories differ from those of all structures in $\mathbb{HYP}_{\mathfrak{M}}$.

I shall show this for the right infinitary languages. It is not so for *finitary* languages. Take any structure $\mathfrak{A} \in \mathbb{HYP}_{\langle \mathfrak{M}, Tr \rangle}$ and consider $Th_{\mathcal{L}}(\mathfrak{A})$, the finitary $\mathcal{L}$-theory of $\mathfrak{A}$. By Lemma (2.3.5) and the transitivity of inductive definitions [40, 1.C.3], we may treat $\mathfrak{A}$ as a Π^1_1 set on $\mathfrak{M}$. As $x \notin Th_{\mathcal{L}}(\mathfrak{A}) \leftrightarrow \neg Sent_{\mathcal{L}}(x) \vee \mathfrak{A} \models \dot{\neg} x$, this shows $Th_{\mathcal{L}}(\mathfrak{A})$ to be in $\mathbb{HYP}_{\mathfrak{M}}$. For strongly acceptable $\mathfrak{M}$, this is enough to show that $Th_{\mathcal{L}}(\mathfrak{A})$ has a model in $\mathbb{HYP}_{\mathfrak{M}}$, by a theorem of Cutland [7]. For a structure $\mathfrak{A}$ with domain M, the same holds for its finitary $\mathcal{K}$-theory. Hence, we do not add structures with new finitary theories when we expand from $\mathbb{HYP}_{\mathfrak{M}}$ to $\mathbb{HYP}_{\langle \mathfrak{M}, Tr \rangle}$.

The main result of this section is that using infinitary theories, we may distinguish some structures in $\mathbb{HYP}_{\langle \mathfrak{M}, Tr \rangle}$ from all those in $\mathbb{HYP}_{\mathfrak{M}}$. I shall show this by showing that we really need all of $\mathcal{L}_{\mathbb{HYP}_{\mathfrak{M}}}$ to define the sets in $\mathbb{HYP}_{\mathfrak{M}}$. We thus find models with distinct infinitary theories all through $\mathbb{HYP}_{\mathfrak{M}}$. This, combined with Theorem (2.5.1), will provide the needed result. This result is the most technical of the paper. Those willing to take my word for it may want to skip to Theorem (2.5.14).

For those interested in the technical details, the approach taken here is to adapt the classical technique of forcing up the ramified analytic hierarchy of Feferman [8]. We may mimic the ramified second-order version of $\mathcal{L}$ by thinking of a second-order quantifier $\exists X^{\alpha}$ of rank α as an infinitary disjunction of formulas of 'lower rank'. Recall that $o(\mathbb{A})$ is the least ordinal not in $\mathbb{A}$, and $O(\mathfrak{M}) = o(\mathbb{HYP}_{\mathfrak{M}})$.

Definition 2.5.2. Let $\mathbb{A}$ be admissible, $o(\mathbb{A}) > \omega$, and let $\mathcal{L} \in \mathbb{A}$. For a fragment $\mathcal{L}_B$, let $\mathrm{DEL}(\mathcal{L}_B)$ be the smallest fragment $\mathcal{L}_C$ such that $\{\bigvee \Phi \mid \Phi \subseteq \mathcal{L}_B \wedge \Phi \in \mathbb{A}\} \subseteq \mathcal{L}_C$. Now let:

$$\begin{aligned} \mathcal{L}_0 &= \mathcal{L}_{\omega\omega} \\ \mathcal{L}_{\alpha+1} &= \mathrm{DEL}(\mathcal{L}_{\alpha}) \\ \mathcal{L}_{\lambda} &= \bigcup_{\sigma<\lambda} \mathcal{L}_{\sigma} \qquad \text{for limit } \lambda \end{aligned}$$

We always have $\bigvee \mathcal{L}_{\alpha} \in \mathcal{L}_{\alpha+1}$. Thus, $\mathcal{L}_{\alpha+1}$ provides something roughly like existential quantification over sets of rank α in the ramified analytic hierarchy. Alternatively, we can see the requirement of membership in $\mathbb{A}$ as building in something like the strength of Δ^1_1-comprehension.

It is clear that the operation $\alpha \mapsto \mathcal{L}_\alpha$ is a Σ-operation in KPU with infinity. Applying Σ-replacement gives:

Lemma 2.5.3. *For admissible* $\mathbb{A}$ *with* $o(\mathbb{A}) > \omega$ *and* $\mathcal{L} \in \mathbb{A}$*, if* $\alpha < o(\mathbb{A})$ *then* $\mathcal{L}_\alpha \in \mathbb{A}$.

In particular, as $O(\mathfrak{M}) > \omega$, for $\sigma < O(\mathfrak{M})$, $\mathcal{L}_\sigma \in \mathbb{HYP}_\mathfrak{M}$. The resolvability of $\mathbb{HYP}_\mathfrak{M}$ allows for something like a converse.

Lemma 2.5.4. *For* $\alpha = O(\mathfrak{M})$*,* $\mathcal{L}_\alpha = \mathcal{L}_{\mathbb{HYP}_\mathfrak{M}}$.

Proof. One direction is trivial. I showed above that for $\sigma < \alpha$, $\mathcal{L}_\sigma \in \mathbb{HYP}_\mathfrak{M}$. By the transitivity of $\mathbb{HYP}_\mathfrak{M}$, we then have $\mathcal{L}_\alpha = \bigcup_{\sigma<\alpha} \mathcal{L}_\sigma \subseteq \mathbb{HYP}_\mathfrak{M}$, and so $\mathcal{L}_\alpha \subseteq \mathcal{L}_{\mathbb{HYP}_\mathfrak{M}}$.

The other direction not much more difficult. $\mathcal{L}_{\mathbb{HYP}_\mathfrak{M}} = \bigcup_{\sigma<\alpha}(\mathrm{L}(M,\sigma) \cap \mathcal{L}_{\infty\omega})$. We have $(\mathrm{L}(M,\omega) \cap \mathcal{L}_{\infty\omega}) \subseteq \mathcal{L}_0 = \mathcal{L}_{\omega\omega}$. Moreover, if $(\mathcal{L}_{\infty\omega} \cap \mathrm{L}(M,\sigma)) \subseteq \mathcal{L}_\gamma$, then $(\mathcal{L}_{\infty\omega} \cap \mathrm{L}(M,\sigma+1)) \subseteq \mathcal{L}_{\gamma+1}$. By induction we thus have $\mathcal{L}_{\mathbb{HYP}_\mathfrak{M}} \subseteq \mathcal{L}_\alpha$. ∎

The same clearly holds for the language with constants $\mathcal{K}$.

There is a simple relation between $\mathcal{K}_{\alpha+1}$ and $\mathbb{HYP}_\mathfrak{M}$. Relying on Lemma (2.3.5) we may restrict our attention to relations on M in $\mathbb{HYP}_\mathfrak{M}$, and make use of the theory of inductive definitions of Moschovakis [40]. Recall that an R-positive φ induces an inductive operator Γ_φ with iterates I^α_φ, least fixed point I_φ, and closure ordinal $\|\Gamma_\varphi\|$.

Theorem 2.5.5. *Let* $S \subseteq M$*,* $S \in \mathbb{HYP}_\mathfrak{M}$*. Then* S *is definable on* $\mathfrak{M}$ *by a* $\mathcal{K}_{\alpha+1}$ *formula for some* $\alpha < O(\mathfrak{M})$.

Proof. Recall that for each $S \in \mathbb{HYP}_\mathfrak{M}$, there is some R-positive formula $\varphi(x_1,\ldots,x_k,y_1,\ldots,y_n)$ and $q_1,\ldots,q_n \in M$ such that $S = I_\varphi(\vec{x},q_1,\ldots,q_n)$ and $\|\Gamma_\varphi\| = \alpha < O(\mathfrak{M})$. Recall also:

$$(\dagger) \qquad I^\alpha_\varphi = \{\langle\vec{x},\vec{y}\rangle \mid \langle\mathfrak{M}, I^{<\alpha}_\varphi\rangle \models \varphi(\vec{x},\vec{y})\}.$$

Now, we may build up the stages of I_φ using $\mathcal{K}_\alpha$. (This is essentially the argument of [1, VI.5.10].) For stage 0, we need to stipulate that R be empty. Let $\varphi(\vec{x},\vec{y},\mathsf{R}/\bot)$ be the result of replacing each occurrence of $\mathsf{R}(t_1,\ldots,t_k)$ in φ by $\bigwedge(t_i \dot{\neq} t_i)$. We can define:

$$\psi_0 = \varphi(\vec{x},\vec{y},\mathsf{R}/\bot)$$
$$\psi_\beta = \varphi(\vec{x},\vec{y},\mathsf{R}/\bigvee_{\gamma<\beta}\psi_\gamma)$$

Using (†), it can easily be proved by induction that

$$(\vec{x},\vec{y}) \in I_\varphi^\beta \leftrightarrow \mathfrak{M} \models \psi_\beta(\vec{x},\vec{y})$$

and

$$(\vec{x},\vec{y}) \in I_\varphi \leftrightarrow \mathfrak{M} \models \bigvee_{\beta<\alpha} \psi_\beta(\vec{x},\vec{y}).$$

Finally, we have $S(\vec{x})$ iff $\mathfrak{M} \models \bigvee_{\beta<\alpha} \psi_\beta(\vec{x}, \dot{\mathsf{q}}_1, \dots \dot{\mathsf{q}}_n)$.

We now must show that $\bigvee_{\beta<\alpha} \psi_\beta(\vec{x}, \dot{\mathsf{q}}_1, \dots \dot{\mathsf{q}}_n)$ is a formula of $\mathcal{K}_{\alpha+1}$. It clearly suffice to show that $\bigvee_{\beta<\alpha} \psi_\beta \in \mathcal{L}_{\alpha+1}$. We proceed by induction.

ψ_0 is clearly in $\mathcal{L}_0 = \mathcal{L}_{\omega\omega}$, and so also in $\mathcal{L}_1$.

Suppose for each $\gamma < \beta \leq \alpha$, $\psi_\gamma \in \mathcal{L}_{\gamma+1}$. We need to show that both ψ_β and $\bigvee_{\gamma<\beta} \psi_\gamma$ are in $\mathcal{L}_{\beta+1}$. Observe that the second implies the first. From the induction hypothesis, $\{\psi_\gamma \mid \gamma < \beta\} \subseteq \mathcal{L}_\sigma$, where $\sigma = \sup\{\gamma+1 \mid \gamma < \beta\}$. $\sigma \leq \beta$, so $\{\psi_\gamma \mid \gamma < \beta\} \subseteq \mathcal{L}_\beta$. Then as $\alpha < O(\mathfrak{M})$, for $\beta \leq \alpha$ $\{\psi_\gamma \mid \gamma < \beta\} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. Hence, from the definition of the $\mathcal{L}_\alpha$-hierarchy, $\bigvee_{\gamma<\beta} \psi_\gamma$ is in $\mathcal{L}_{\beta+1}$.

By induction, we thus have $\bigvee_{\beta<\alpha} \psi_\beta \in \mathcal{L}_{\alpha+1}$. ■

I shall now show that for any $\alpha < O(\mathfrak{M})$, there is an S hyperelementary on $\mathfrak{M}$ that cannot be defined in $\mathcal{K}_{\alpha+1}$, and hence cannot be projected from a fixed point I_φ^α. This will show that we find non-trivial definability properties throughout all of $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$.

This will be done by an analog of Cohen forcing up the ramified analytic hierarchy, with $\mathcal{K}_\alpha$ playing the role of the ramified analytic hierarchy. The classical version of these ideas was developed by Feferman [8]. My presentation in part borrows from Sacks [47]. Let S be a new predicate symbol, and let $\mathcal{K}_\alpha(\mathsf{S})$ be the smallest fragment containing both $\mathcal{K}_\alpha$ and S. Thus, S figures into no infinitary combinations. (In the ramified second-order case, we add a new constant to a language which already has (ranked and unranked) second-order variables.)

Define a *forcing condition* to be a finite consistent sets of sentences of the form $\mathsf{S}(\dot{\mathsf{m}})$ or $\neg\mathsf{S}(\dot{\mathsf{m}})$. The forcing conditions form a partial order $\langle \mathcal{P}, \geq \rangle$ under the inclusion relation $\subseteq$. The notion of consistency for such simple sets is entirely syntactic: p is consistent if for no $\dot{\mathsf{m}}$ is both $\mathsf{S}(\dot{\mathsf{m}})$ and $\neg\mathsf{S}(\dot{\mathsf{m}})$ in p. Hence, the partial order $\langle \mathcal{P}, \geq \rangle \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$.

A forcing relation is defined as usual.

Definition 2.5.6. The *forcing relation* between conditions and sentences of $\mathcal{K}_\alpha(\mathsf{S})$ is defined as follows.

1. $p \Vdash \mathsf{R}_i(\dot{\mathsf{m}}_1, \dots, \dot{\mathsf{m}}_n)$ iff $\mathfrak{M} \models \mathsf{R}_i(\dot{\mathsf{m}}_1, \dots, \dot{\mathsf{m}}_n)$, for atomic $\mathsf{R}_i(\dot{\mathsf{m}}_1, \dots, \dot{\mathsf{m}}_n)$ in $\mathcal{K}$.

2. $p \Vdash \mathsf{S}(\dot{\mathsf{m}})$ iff $\mathsf{S}(\dot{\mathsf{m}}) \in p$.

3. $p \Vdash \varphi \vee \psi$ iff $p \Vdash \varphi$ or $p \Vdash \psi$.

4. $p \Vdash \neg\varphi$ iff for all $q \leq p$, $q \nVdash \varphi$.

5. $p \Vdash \exists v\, \varphi$ iff for some m, $p \Vdash \varphi(\dot{\mathsf{m}})$.

6. $p \Vdash \bigvee \Phi$ iff for some $\varphi \in \Phi$, $p \Vdash \varphi$, where $\bigvee \Phi \in \mathrm{DEL}(\mathcal{K}_\beta)$ for some $\beta < \alpha$.

As $\mathfrak{M}$ will remain fixed throughout, let us write $X \models \varphi$ for $\langle \mathfrak{M}, X \rangle \models \varphi$. We may say $S \in p$ if $S \models \bigwedge p$, i.e. S satisfies each member of p.

In some applications of forcing in recursion theory, such as Hinman [24], it is useful to define a restricted notion of genericity. We do the same thing here.

Definition 2.5.7. A predicate S on M is *α-generic* if for each sentence φ of $\mathcal{K}_\alpha(\mathsf{S})$ there is a condition p such that $S \in p$ and $p \Vdash \varphi$ or $p \Vdash \neg\varphi$.

Many of the usual facts about generics and forcing carry over to our setting. The usual definability of forcing shows in this setting that for $\alpha < O(\mathfrak{M})$, the forcing relation for $\mathcal{K}_\alpha(\mathsf{S})$ is in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. Most importantly, we have versions of the truth lemma and the existence of generics.

Lemma 2.5.8. *If S is α-generic then for each $\mathcal{K}_\alpha(\mathsf{S})$-sentence φ,*

$$S \models \varphi \quad \textit{iff} \quad \exists p (S \in p \wedge p \Vdash \varphi).$$

The proof is nearly the same as the one for ramified analysis, as in [47, IV.3.5]. Modifying familiar proofs also gives us the following.

Lemma 2.5.9. *For $\mathfrak{M}$ strongly acceptable and $\alpha < O(\mathfrak{M})$, there are α-generics in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$.*

There is an important connection between definability and α-genericity.

Theorem 2.5.10. *If S is α-generic, then S cannot be defined by a $\mathcal{K}_\alpha$-formula.*

To prove this, it will be helpful to have a technical lemma.

Definition 2.5.11. Fix some $d \in M$.

1. For a condition p, define a condition $p^\star$ which disagrees with p on d by:

$$\mathsf{S}(\dot{\mathsf{m}}) \in p^\star \quad \text{iff} \quad \begin{cases} m \neq d & \text{and} \quad \mathsf{S}(\dot{\mathsf{m}}) \in p \\ m = d & \text{and} \quad \neg\mathsf{S}(\dot{\mathsf{m}}) \in p \end{cases}$$

$$\neg\mathsf{S}(\dot{\mathsf{m}}) \in p^\star \quad \text{iff} \quad \begin{cases} m \neq d & \text{and} \quad \neg\mathsf{S}(\dot{\mathsf{m}}) \in p \\ m = d & \text{and} \quad \mathsf{S}(\dot{\mathsf{m}}) \in p \end{cases}$$

2. For a formula φ, $\varphi^\star$ is the formula that results from substituting $[(t \dot{\neq} \dot{\mathsf{d}} \wedge \mathsf{S}(t)) \vee (t \doteq \dot{\mathsf{d}} \wedge \neg\mathsf{S}(t))]$ for each occurrence of $\mathsf{S}(t)$.

Lemma 2.5.12 (Feferman). *$p \Vdash \varphi$ iff $p^\star \Vdash \varphi^\star$.*

The proof is essentially that given in Feferman [8]. We may now give:

Proof of (2.5.10). For contradiction, suppose S is α-generic and S is defined by a $\mathcal{K}_\alpha$-formula φ. We then have $S \models \forall v(\varphi(v) \leftrightarrow \mathsf{S}(v))$. As $\forall v(\varphi(v) \leftrightarrow \mathsf{S}(v))$ is a $\mathcal{K}_\alpha(\mathsf{S})$-sentence, and S is generic, by Lemma (2.5.8), there must be some p such that $S \in p$ and $p \Vdash \forall v(\varphi(v) \leftrightarrow \mathsf{S}(v))$.

Fix such a p. p is a finite set of conditions, so there must be some d such that neither $\mathsf{S}(\dot{\mathsf{d}})$ nor $\neg\mathsf{S}(\dot{\mathsf{d}})$ is in p. Fix some such d. Relative to this d, observe that $p^\star = p$. Hence, by the last Lemma (2.5.12), $p \Vdash [\forall v(\varphi(v) \leftrightarrow \mathsf{S}(v))]^\star$. By our choice of φ, this means $p \Vdash \forall v(\varphi(v) \leftrightarrow (\mathsf{S}(v))^\star)$. By Lemma (2.5.8), we thus have $S \models \forall v(\varphi(v) \leftrightarrow (\mathsf{S}(v))^\star)$. For the d instance, unpacking the definition of $\star$, this implies $S \models \varphi(\dot{\mathsf{d}}) \leftrightarrow \neg\mathsf{S}(\dot{\mathsf{d}})$. This contradicts our assumption that φ defines S, so we are done. ■

Combining Theorem (2.5.10) with (the proof of) Theorem (2.5.5) we see:

Corollary 2.5.13. *If $S \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ is α-generic and $S = I_\varphi(m_1, \ldots, m_k)$, then $\|\Gamma_\varphi\| \geq \alpha$.*

These results show that for each level α of $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$, we can find sets that we need to go as high as α to build, and we need as much as $\mathcal{K}_\alpha$ to describe. We thus really do use all of $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ to model the expressive resources provided by context $\mathfrak{M}$. As a result, when we expand $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ to $\mathbb{H}\mathrm{YP}_{\langle \mathfrak{M}, Tr\rangle}$, we get genuinely new structures.

Theorem 2.5.14. *Let $\alpha = O(\mathfrak{M})$. Let P be inductive non-hyperelementary on $\mathfrak{M}$. There is a structure $\mathfrak{A}$ in $\mathbb{H}\mathrm{YP}_{\langle \mathfrak{M}, P\rangle}$ such that no structure in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ models $Th_{\mathcal{K}_\alpha}(\mathfrak{A})$ ($= Th_{\mathcal{K}_{\mathbb{H}\mathrm{YP}_{\mathfrak{M}}}}(\mathfrak{A}) \in \mathbb{H}\mathrm{YP}_{\langle \mathfrak{M}, P\rangle}$). Hence, no structure in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ models $Th_{\mathcal{K}_{\mathbb{H}\mathrm{YP}_{\langle \mathfrak{M}, P\rangle}}}(\mathfrak{A})$ as well.*

Proof. First, recall that $\alpha \in \mathbb{H}\mathrm{YP}_{\langle\mathfrak{M},P\rangle}$. Hence there is an α-generic S in $\mathbb{H}\mathrm{YP}_{\langle\mathfrak{M},P\rangle}$, and a structure $\langle\mathfrak{M}, S\rangle \in \mathbb{H}\mathrm{YP}_{\langle\mathfrak{M},P\rangle}$. From Corollary (2.5.13), we know that $S \notin \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$, and so $\langle\mathfrak{M}, S\rangle \notin \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. Moreover,

$$Th_{\mathcal{K}_\alpha(\mathsf{S})}(\langle\mathfrak{M}, S\rangle) = \{\varphi \in \mathcal{K}_\alpha(\mathsf{S}) \mid \exists p(S \in p \wedge p \Vdash \varphi)\}$$

by Lemma (2.5.8).

Suppose $(\langle\mathfrak{B}, Q\rangle, b_m)_{m\in M} \models Th_{\mathcal{K}_\alpha(\mathsf{S})}(\langle\mathfrak{M}, S\rangle)$ and $(\langle\mathfrak{B}, Q\rangle, b_m)_{m\in M} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. Let $Q' = \{m \mid (\langle\mathfrak{B}, Q\rangle, b_m)_{m\in M} \models \mathsf{S}(\dot{\mathsf{m}})\}$. Q' is in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ if Q is, as satisfaction is Δ in KPU. Thus the $\mathcal{K}(\mathsf{S})$-structure $\langle\mathfrak{M}, Q'\rangle$ is in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. I claim Q' is α-generic. As S is α-generic, for any $\mathcal{K}_\alpha(\mathsf{S})$-sentence φ, there is condition p such that $S \in p$ and p decides φ. We need to show that $Q' \in p$, i.e. $Q' \models \bigwedge p$. By the definition of Q', we see that $Q' \models \pm\mathsf{S}(\dot{\mathsf{m}})$ iff $(\langle\mathfrak{B}, Q\rangle, b_m) \models \pm\mathsf{S}(\dot{\mathsf{m}})$. But $\langle\mathfrak{B}, Q\rangle$ is chosen to agree with $\langle\mathfrak{M}, S\rangle$ on $\mathcal{K}_\alpha(\mathsf{S})$, so $(\langle\mathfrak{B}, Q\rangle, b_m) \models \bigwedge p$ iff $S \models \bigwedge p$. Thus $Q' \models \bigwedge p$ iff $S \models \bigwedge p$, and so $Q' \in p$.

Finally, we observe that as Q' is α-generic, by Corollary (2.5.13) Q' cannot be in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. This is a contradiction, so the result is proved for $\mathcal{K}(\mathsf{S})$.

To see that this works for $\mathcal{K}$ as well, let R be some predicate symbol in $\mathcal{K}$ interpreted by R in $\mathfrak{M}$. (For simplicity, we may suppose R to be unary.) Using the acceptability of $\mathfrak{M}$, we may define a new predicate $\bar{R} = \{x \mid \exists y[(x = \langle 0, y\rangle \wedge R(y)) \vee (x = \langle 1, y\rangle \wedge S(y))]\}$. Let $\overline{\mathfrak{M}}$ be $\mathfrak{M}$ with $\bar{R}$ substituted for R. Clearly $\overline{\mathfrak{M}}$ is in $\mathbb{H}\mathrm{YP}_{\langle\mathfrak{M},P\rangle}$ but not in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$.

We now proceed pretty much as before. We may introduce a *defined symbol* $\dot{\mathsf{S}}$ by $\dot{\mathsf{S}}(v) \leftrightarrow \exists w(\mathsf{R}(w) \wedge w = \langle 1, v\rangle)$. Suppose there is some $\mathfrak{B} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ such that $(\mathfrak{B}, b_m)_{m\in M} \models Th_{\mathcal{K}_\alpha}(\overline{\mathfrak{M}})$. We define $\bar{Q}$ much like Q', using the defined $\dot{\mathsf{S}}$. $\bar{Q} = \{m \mid (\mathfrak{B}, b_m) \models \dot{\mathsf{S}}(\dot{\mathsf{m}})\}$. If $\mathfrak{B}$ is in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$, $\bar{Q}$ is as well. However, an easy calculation, similar to the one above, shows that this implies S is in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. As S is α-generic, this is a contradiction, so we are done. ∎

This solves the expansion problem. There are more structures—more truth conditions—in $\mathbb{H}\mathrm{YP}_{\langle\mathfrak{M},Tr\rangle}$ than there are in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$; more structures that can be distinguished by resources available in $\mathbb{H}\mathrm{YP}_{\langle\mathfrak{M},Tr\rangle}$.

2.6 Infinitary Languages and Semantic Relations

There is one loose end that can now be (partially) tied up. My formal model of the expressive resources speakers can bring to bear in a context $\mathfrak{M}$ has been $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. In particular, in Theorem (2.5.14) I worked with the language

$\mathcal{K}_{\mathbb{H}\mathrm{YP}_{\mathfrak{M}}}$. But as I mentioned, we expect the resources speakers can bring to bear to be given in part by the language they use, which is much more directly modeled by the language $\mathcal{K}^+$ containing a semantic relation. In Section (1.2.2) I suggested that we can see a match-up between the language with semantic relations and the right sort of infinitary language, and have been proceeding on that assumption.

We can now make this match-up more precise.

Theorem 2.6.1. *Distinct structures in* $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ *can be distinguished by sentences of* $\mathcal{K}^+$.

Proof. Let $\mathfrak{A}$ and $\mathfrak{B}$ be distinct structures in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. From Corollary (2.3.4), there is some sentence ψ of $\mathcal{K}_{\mathbb{H}\mathrm{YP}_{\mathfrak{M}}}$, indeed of $\mathcal{L}_{\mathbb{H}\mathrm{YP}_{\mathfrak{M}}}$, such that $\mathfrak{A} \models \psi$ while $\mathfrak{B} \models \neg\psi$. We need to find a sentence ψ' of $\mathcal{K}^+$ that expresses this. ψ' is built up inductively. For atomic ψ, we want $\psi' = \mathsf{Tr}(\pi(\dot{\mathfrak{A}}), \ulcorner\psi\urcorner) \wedge \mathsf{Tr}(\pi(\dot{\mathfrak{B}}), \ulcorner\neg\psi\urcorner)$. The interesting induction step is for $\psi = \bigvee\Phi$. As $\psi \in \mathcal{L}_{\mathbb{H}\mathrm{YP}_{\mathfrak{M}}}$, Φ is in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. Now, relying upon Theorem (2.4.6), we may invoke a classic result (as in [38, 5.21]), that for $S \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$, S is fully definable by a formula of $\mathcal{K}^+$ (i.e. there is some formula γ such that $m \in S$ iff $TM(\ulcorner\gamma(\dot{\mathsf{m}})\urcorner)$ iff $\neg TM(\ulcorner\neg\gamma(\dot{\mathsf{m}})\urcorner)$). Hence, there is a formula γ' of $\mathcal{K}^+$ that defines the set $\{\ulcorner\varphi'\urcorner \mid \varphi \in \Phi\}$ on $\langle\mathfrak{M}, \mathit{Tr}\rangle$. We then let $\psi' = \exists v(\gamma'(v) \wedge (\mathsf{Tr}(\pi(\dot{\mathfrak{A}}), v) \wedge \mathsf{Tr}(\pi(\dot{\mathfrak{B}}), \dot{\neg}v)))$. We always have $\langle\mathfrak{M}, \mathit{Tr}\rangle \models \psi'$. ■

We are thus safe in saying that indeed speakers in a context $\mathfrak{M}$ can distinguish the structures in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ using $\mathcal{K}^+$. As the relation Tr relies on the domain of structures in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ for its interpretation, it displays the same extraordinary context dependence as $\exists p$. The theorem shows that relative to a given context $\mathfrak{M}$, the expressive resources of $\mathcal{K}_{\mathbb{H}\mathrm{YP}_{\mathfrak{M}}}$ are no stronger than those of $\mathcal{K}^+$. We will see in Remark (3.3.12) of Section (3.3) that this relation goes the other way as well. But the direction established by Theorem (2.6.1) is already enough to justify the use I have made of infinitary languages, for it guarantees that having found new structures in $\mathbb{H}\mathrm{YP}_{\langle\mathfrak{M}, \mathit{Tr}\rangle}$ by lights of $\mathcal{K}_{\mathbb{H}\mathrm{YP}_{\langle\mathfrak{M}, \mathit{Tr}\rangle}}$ is enough to have found new structures by lights of $\mathcal{K}^+$ in context $\langle\mathfrak{M}, \mathit{Tr}\rangle$.

2.7 The Formal Model of Extraordinary Context Dependence

We have now gone a long way towards making the behavior of the Liar much more tractable. I suggest that the expansion problem has been resolved; and with it, the basic mysteries of extraordinary context dependence.

There were two crucial things that needed to be explained, as I mentioned in Section (1.1.4). First, we needed to explain how there is any context shift in the Liar inference, between contexts (A) and (B) in section (1.1.3). We have seen that part of context is described by a salience structure, and that in this crucial step, the salience structure expands by adding a semantic relation, given by Tr in the formal model. The formal model captures this step in the shift from $\mathfrak{M}$ to $\langle \mathfrak{M}, Tr \rangle$. Second, we needed to explain how this can effect an expansion of the background domain of truth conditions. We have seen that the truth conditions available for constructing propositions relative to a context are constrained by the expressive resources speakers can bring to bear in that context. I have suggested this is well-modeled by the structure $\mathbb{HYP}_{\mathfrak{M}}$. The truth conditions available are modeled by the structures in $\mathbb{HYP}_{\mathfrak{M}}$. The results of Section (2.3) support this. As we have seen in Theorem (2.5.14), there are genuinely more structures in $\mathbb{HYP}_{\langle \mathfrak{M}, Tr \rangle}$ than there were in $\mathbb{HYP}_{\mathfrak{M}}$. Thus, the context shift really does lead to an expansion of the background domain of truth conditions.

In Section (1.1.5) I compared this sort of expansiveness with the behavior of situations. I think we can see in the results of Section (2.5) a more satisfying theory. The structures in $\mathbb{HYP}_{\langle \mathfrak{M}, Tr \rangle}$ are still $\mathcal{L}$-structures, just as those in the initial $\mathbb{HYP}_{\mathfrak{M}}$ were. But they are structures that code up much more complex relations than those in $\mathbb{HYP}_{\mathfrak{M}}$, reflecting the expanded expressive resources provided by the expanded context $\langle \mathfrak{M}, Tr \rangle$. Rather than simply throwing extra semantic relations onto structured truth conditions, which I rejected in Section (1.1.5), we can now see that a semantic relation does genuinely become salient in the Liar inference. This expands the salience structure, and that in turn allow for greater *complexity* in truth conditions. The result is expansion. I shall show in Section (3.3) how this greater complexity can be used to reconstruct some more situation-like structure in truth conditions, but for the moment, it should be stressed that it is complexity that is doing the important work.

This expansiveness is the basis of the context dependence in the Liar, as I argued in Section (1.1.3). The Liar sentence contains a tacit quantifier over the domain of truth conditions, or the domain of propositions that can be build from it. Though this domain is not in an ordinary way contextually restricted, we have seen that it can shift as context shifts. Hence, we have a phenomenon of extraordinary context dependence. The particular dependence we have observed opens the way for resolving the Liar paradox. Relative to the initial context, with the smaller domain of truth conditions, there is no proposition for the Liar sentence to express. Relative to the expanded domain of the shifted context, there is. Hence, our two apparently paradoxical conclusions at (A) and (B) need not be paradoxical at all.

3 Hierarchy

So far, we have addressed what I take to be the basic phenomenon underlying the Liar paradox. In investigating extraordinary context dependence, we have seen how the Liar sentence can be context dependent. We have seen that there is a context shift in the Liar inference, and that it leads to an expansion of the background domain of truth conditions to which the Liar sentence is sensitive. This does the heavy lifting of resolving the paradox. As there is the right kind of context shift, and the right kind of context dependence, the apparently intractable pair of conclusions we reach in the Liar inference are no longer intractable. They can both be correct, as the very same sentence can in the first context be unable to express a proposition, and then express one in the second.

Though the heavy lifting has been done, the detail work has not. Bitter experience with the paradox has shown that the devil really can be lurking in the details. Though in solving the resolution problem, I have shown that it is coherent for the Liar sentence to express a proposition in some contexts, I have not yet fully explained what that proposition is, or how the Liar sentence expresses it. Indeed it is not yet clear why, in spite of the expansion of the domain of truth conditions, the Liar sentence does not just continue in its previous semantic status.

In this part, I shall turn to the detail work. I shall argue that the hierarchical structure implicit in the context-based solution to the expansion problem I have offered helps us to address these remaining issues. In the first Section (3.1), I shall show how a hierarchical structure emerges from the phenomenon of extraordinary context dependence. I shall then show that this hierarchy has a role to play in answering the remaining questions about the Liar inference in Section (3.2). I shall show in detail how it plays this role in Section (3.3), by further developing the formal model of Part (2). With the tools provided by this development in hand, I shall return in Section (3.4) to the Liar inference as originally presented in Section (1.1.3), and show in detail how the paradox is avoided. I shall conclude the paper by examining the nature of the hierarchy that emerges from context dependence in Section (3.5), and finally by showing how it relates to an important more general idea about hierarchies in Section (3.6).

3.1 The Emergent Hierarchy

Out of the resolution of the expansion problem, a hierarchy emerges. Fundamentally, we have a hierarchy of contexts and domains of truth conditions determined by those contexts. For any context $\mathfrak{M}$, we can find a distinct context $\langle \mathfrak{M}, Tr \rangle$. The domains associated with these contexts form a hierarchical structure, as we know $\{\mathfrak{A} \mid \mathfrak{A} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}\} \subsetneq \{\mathfrak{A} \mid \mathfrak{A} \in \mathbb{H}\mathrm{YP}_{\langle \mathfrak{M}, Tr \rangle}\}$. If we like, we could count this as two hierarchies, one of contexts and one of domains of truth conditions. But the two hierarchies are intimately related, and generated by the same underlying phenomenon of extraordinary context dependence. I think it is better to describe what we have as one basic underlying hierarchy with several aspects. Indeed, we can identify a third aspect of this basic hierarchy, as the distinct domains of truth conditions induce distinct semantic relations $Tr^{\mathfrak{M}}$ and $Tr^{\langle \mathfrak{M}, Tr \rangle}$. Each of these aspects—each hierarchical structure—is open-ended. The step from $\mathfrak{M}$ to $\langle \mathfrak{M}, Tr \rangle$ can be repeated, giving a context $\langle \mathfrak{M}, Tr^{\mathfrak{M}}, Tr^{\langle \mathfrak{M}, Tr \rangle} \rangle$, and so on. We likewise have open-ended sequences of expanded domains of truth conditions, and semantic relations. It is not obvious that these need to be linearly ordered. But as I shall discuss in Section (3.5), the sequence of domains of truth conditions, and hence the sequence of semantic relations, are predicted to be linearly ordered by the mathematics I have used to analyze the situation.

As I promised, this hierarchy, in all its aspects, has emerged from the underlying phenomenon of extraordinary context dependence. We worked out what aspects of context are involved in the Liar, and how the Liar sentence is dependent upon them, and found the phenomenon of domain expansion. This can be repeated, to induce a fundamental hierarchical structure with several different aspects. It is clearly a more liberal hierarchy than the one originally suggested by Tarski. It does not place any syntactic restrictions on the semantic relations, and the Kripke-inspired semantics of Tr allows for a great deal of semantic closure. But it does fundamentally divide basic semantic apparatus into levels, and so is essentially hierarchical.

To some extent, the hierarchy may appear epiphenomenal, as emergent structure often is. But we will see in the following sections that is is useful in addressing the yet-unanswered questions about the propositions expressed by the Liar sentence.[22]

[22]Most context-based approaches to the paradox, including those of Barwise and Etchemendy [2], Burge [3], Gaifman [10, 11], Koons [30], and Parsons [43], wind up with some aspects of hierarchy, though these authors differ on the nature of the hierarchy and the amount of emphasis they place upon it. A more difficult case to classify is Simmons [48, 49]. He comes to some conclusions that seem to me to be hierarchical, but explicitly

3.2 Reasoning Across Contexts

The role for the hierarchy derives from one more consideration of context dependence. Generally, the presence of context dependence makes reasoning—or rather, stating ones reasoning—a more difficult task. As sentences can shift what they express across contexts, the correctness of an inferential step across context can be a delicate matter. We see this in a dramatic way in the Liar inference.

First, consider a simpler example. We are at the aquarium. Demonstrating a whale, you say 'That is a whale'. A few moments later, demonstrating an eel, you cannot correctly draw the conclusion expressed by 'That is a mammal'. The reason is not anything mysterious about the entailment from being a whale to being a mammal. That still holds. The problem is simply one of what is expressed by the second utterance 'That is a mammal', and particularly that the referent of 'that' has changed with the context. Especially in reporting reasoning across contexts, speakers have to be careful to express their claims correctly. In this case, the correct report for the second step is something like 'The previously demonstrated animal is a mammal'.

Here we see speakers making a claim that depends upon a prior claim. When they do, they need to make sure they express the right content, the content correctly related to the prior claim.

With this in mind, let us look back once more at the crucial steps in the Liar inference:

(A) $$\neg\exists p(\mathsf{Exp}(\ulcorner l \urcorner, p).$$

(B) $$l.$$

In the first, we conclude l does not express a proposition. We then reasoned by logic, that it therefore does not express a true proposition. This was the assertion of l at (B). We could go on to conclude l expresses a true proposition after all.

As with the 'whale' example, we need to be careful about this sort of reasoning across the context shift from (A) to (B). Now, there is no real worry about the referent of l, which is the same sentence in both contexts. But we do have to worry about the domain of $\exists p$. The trivial step from

$$\neg\exists p(\mathsf{Exp}(\ulcorner l \urcorner, p)$$

to

$$\neg\exists p(\mathsf{Exp}(\ulcorner l \urcorner, p) \wedge \mathsf{Tr}(p))$$

rejects hierarchies *per se*. I suggest below that perhaps some of his objections are targeted at a more traditional, Tarskian hierarchy than the one I propose here.

is only trivial if the domain of $\exists p$ has not expanded during it. And of course, it has. Just as with the 'whale' case, where the correct step was reported by 'the previously demonstrated mammal', here the trivial step is one which would roughly be reported as 'As used in context (A), l does not express a true proposition'. We thus find ourselves needing to make reference to the semantics potentials of sentences in context, in order to clearly state the logically trivial inference. To introduce some terminology, I shall say we make reference to the *semantics of a context* when we make reference to the semantic properties of sentences in the context. To state the trivial inference properly, we need to make reference to the semantics of the initial context (A).

It is tempting to conclude that in the Liar case, this is easy. After all, at (B) the context looks like $\langle \mathfrak{M}, Tr \rangle$, so speakers have readily salient and available the relation Tr. With it they can recover the predicate Str (and indeed $\mathbb{H}\mathrm{YP}_{\mathfrak{M}} \in \mathbb{H}\mathrm{YP}_{\langle \mathfrak{M}, Tr \rangle}$), so they can easily reconstruct the domain of $\exists p$ in the initial context (A). With that, the can describe the semantics of (A) well enough to express the right proposition.

This is to some extent right, but it under-estimates the task speakers face. In the salience structure $\langle \mathfrak{M}, Tr \rangle$, Tr is just another salient relation. I showed in Part (2) that because of its complexity relative to $\mathfrak{M}$, its presence has important consequences. But as an element of the salience structure, it is still just one of many salient relations. In particular, it is not singled out as a semantic relation of the previous context. Likewise Str (or $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$) is not singled out as giving the domain of truth conditions of the previous context. Thus, the path from the availability of these relations to the reconstruction of the semantics of the previous context is not so direct.

Let us return to the'whale' example once more. There, the task speakers face is not simply to re-identify an animal across time. Rather, they need to identify something in the second context as the referent of an utterance in the first context. They need to identify an object with respect to the semantics of the previous context. Likewise, in the Liar inference, speakers need not just identify the relation Tr; they must identify it *as* the semantic relation of that context.

How do they do this? I shall present a formal model continuing the work of Part (2) below. But first, let me informally sketch what will be involved. As in the 'whale' case, speakers need to identify features of the prior context, and on the basis of them work out how to express aspects of its semantics from their new context. First, they identify the prior context itself, or at least its salience structure. From $\langle \mathfrak{M}, Tr \rangle$, it is easy enough to identify $\mathfrak{M}$. They then start rebuilding the semantic relation $Tr^{\mathfrak{M}}$ of $\mathfrak{M}$. They rebuild it within $\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}$, but as the semantic relation of $\mathfrak{M}$. I shall suggest they may do this piecewise. With the resources of $\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}$, they can come to see each truth condition in $\mathbb{HYP}_{\mathfrak{M}}$ as implicitly coding up its contribution to the semantics of $\mathfrak{M}$. In $\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}$, they can extract this information, and use it to reconstruct the semantic relation Tr of $\mathfrak{M}$. This displays Tr as the semantics of $\mathfrak{M}$ within $\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}$.

The emergent hierarchy I discussed in the previous section is important for enabling speakers to do this. They need the expanded resources of the new context to reconstruct the semantics of the prior context. But we will also see that in reconstructing the semantic relations of the prior context, speakers are led to a more traditional aspect of hierarchy. Within the new context they have available multiple semantic relations: the proper one of the context, and the one that reconstructs the semantics of the prior context.

3.3 Internal Truth Relations

We now need to see some more details of how the process of reconstructing the semantic relation of the prior context may proceed. Continuing the project of Part (2), my goal here is to work a simplified example in detail, by spelling it out in the terms of the formal model I proposed there. As we will now be looking at the truth relations of several contexts at once, some notation is in order. Write $Tr^{\mathfrak{M}}$ for the truth relation relative to context $\mathfrak{M}$. As usual, we think of $\mathfrak{M}$ as the context, and $\mathbb{HYP}_{\mathfrak{M}}$ as providing the resources that can be brought to bear in that context.

The crucial informal idea is that from the perspective of the expanded context, i.e. in $\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}$, some structures implicitly give information about the semantics of the prior context, which can be extracted. This is spelled out in the following definition. By virtue of Corollary (2.3.4), each structure in $\mathbb{HYP}_{\mathfrak{M}}$ can be thought of as given by its $\mathcal{K}_{\mathbb{HYP}_{\mathfrak{M}}}$-theory. A structure implicitly codes up information about a set S relative to $\mathbb{HYP}_{\mathfrak{M}}$ if we can move from its $\mathcal{K}_{\mathbb{HYP}_{\mathfrak{M}}}$-theory to S within $\mathbb{HYP}_{\mathfrak{M}}$.

Definition 3.3.1. Let $\mathfrak{A}$ be in $\mathbb{HYP}_{\mathfrak{M}}$. $\mathfrak{A}$ *articulates* a set S in $\mathbb{HYP}_{\mathfrak{M}}$ if there is a subset of its $\mathcal{K}_{\mathbb{HYP}_{\mathfrak{M}}}$-theory in $\mathbb{HYP}_{\mathfrak{M}}$ and a mapping in $\mathbb{HYP}_{\mathfrak{M}}$ of that subset onto S.

We want to say that structures in $\mathbb{HYP}_{\mathfrak{M}}$ articulate their contributions to the parameterized truth relation $Tr^{\mathfrak{M}}$ relative to the expanded context $\langle\mathfrak{M}, Tr\rangle$. To see the contribution of a structure to Tr, we simply restrict our attention to that structure in the definition.

Definition 3.3.2. Define $\mathfrak{A} \upharpoonright Tr$, the $\mathfrak{A}$-*section of* Tr, by:

$$\mathfrak{A} \upharpoonright Tr_0 = \{\langle x, y\rangle \mid x = \pi(\mathfrak{A}) \wedge Sent(y) \wedge \langle x, \emptyset\rangle \models y^*\}$$
$$\mathfrak{A} \upharpoonright Tr_{\sigma+1} = \{\langle x, y\rangle \mid x = \pi(\mathfrak{A}) \wedge Sent(y) \wedge \langle x, \mathfrak{A} \upharpoonright Tr_\sigma\rangle \models y^*\}$$
$$\mathfrak{A} \upharpoonright Tr = \cup_\sigma \mathfrak{A} \upharpoonright Tr_\sigma$$

We know that $\|\mathfrak{A} \upharpoonright Tr\| \leq O(\mathfrak{M})$ for each $\mathfrak{A} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$.

For sentences in an $\mathfrak{A}$-section, the only structure term that makes any difference is $\dot{\pi(\mathfrak{A})}$. Thus, quantification into the first argument place of Tr is essentially the same as having $\dot{\pi(\mathfrak{A})}$ in that position. For instance, $\langle \pi(\mathfrak{A}), \ulcorner \exists w \mathsf{Tr}(w, \ulcorner \dot{\mathsf{m}} \doteq \dot{\mathsf{m}}\urcorner)\urcorner\rangle \in \mathfrak{A} \upharpoonright Tr_1$, but only because $\langle \pi(\mathfrak{A}), \ulcorner \dot{\mathsf{m}} \doteq \dot{\mathsf{m}}\urcorner\rangle \in \mathfrak{A} \upharpoonright Tr_0$.

We can thus essentially think of an $\mathfrak{A}$-section as involving only sentences containing a *unary* predicate $\mathsf{T}_{\mathfrak{A}}$. We may make this precise by considering a language $\mathcal{K}^{\mathsf{T}_{\mathfrak{A}}}$ which extends $\mathcal{K}$ by adding the symbol $\mathsf{T}_{\mathfrak{A}}$. For $\mathcal{K}^{\mathsf{T}_{\mathfrak{A}}}$, we can carry out the Kripke construction in its usual form.

Definition 3.3.3.

$$T_{\mathfrak{A},0} = \{\ulcorner \varphi \urcorner \mid Sent_{\mathcal{K}^{\mathsf{T}_{\mathfrak{A}}}}(\varphi) \wedge \langle \mathfrak{A}, \emptyset\rangle \models \varphi^*\}.$$
$$T_{\mathfrak{A},\sigma+1} = \{\ulcorner \varphi \urcorner \mid Sent_{\mathcal{K}^{\mathsf{T}_{\mathfrak{A}}}}(\varphi) \wedge \langle \mathfrak{A}, T_{\mathfrak{A},\sigma}\rangle \models \varphi^*\}$$
$$T_{\mathfrak{A}} = \bigcup_\sigma T_{\mathfrak{A},\sigma}$$

$T_{\mathfrak{A}}$ is essentially the $\mathfrak{A}$-section $\mathfrak{A} \upharpoonright Tr$ of Tr, cutting out essentially vacuous quantification into the structure position of Tr. To make this more precise, for a fixed $\mathfrak{A}$, we may define a map $\Upsilon\colon \mathcal{K}^+ \to \mathcal{K}^{\mathsf{T}_{\mathfrak{A}}}$ which replaces each occurrence of $\mathsf{Tr}(x, y)$ with $\mathsf{T}_{\mathfrak{A}}(y)$. Such an Υ may be extended to a map on Tr by setting $\bar{\Upsilon}(\langle \pi(\mathfrak{B}), \ulcorner \varphi \urcorner\rangle) = \ulcorner \Upsilon(\varphi)\urcorner$. $\bar{\Upsilon}$ is truth preserving on the $\mathfrak{A}$-section of Tr, and matches up elements of $\mathfrak{A} \upharpoonright Tr$ with those of $T_{\mathfrak{A}}$. It does so in a way that preserves levels, as the following lemma shows.

Lemma 3.3.4. $\bar{\Upsilon}$ *maps* $\mathfrak{A} \upharpoonright Tr_\alpha$ onto $T_{\mathfrak{A},\alpha}$.

Proof. The proof is by induction on α and then on formulas. For the latter, take the base case to be atomic and negated atomic sentences, treating $\neg\mathsf{Tr}$ separately, and observe that we may assume for sentences appearing in $\mathfrak{A} \upharpoonright Tr$ that:

$$\Upsilon(\mathsf{Tr}(x, \ulcorner \dot{\varphi} \urcorner)) = \mathsf{T}_{\mathfrak{A}}(\ulcorner \dot{\Upsilon(\varphi)}\urcorner).$$

To see that the map is onto, define a map Υ^{-1} from formulas of $\mathcal{K}^{\mathsf{T}_{\mathfrak{A}}}$ to $\mathcal{K}^+$ by replacing each occurrence of $\mathsf{T}_{\mathfrak{A}}(x)$ by $\mathsf{Tr}(\dot{\pi(\mathfrak{A})}, x)$. ■

This shows that we can think of each $T_{\mathfrak{A}}$ as a kind of cleaned-up version of the corresponding $\mathfrak{A}$-section of Tr. $T_{\mathfrak{A}}$ is much nicer to work with than Tr, as Definition (3.3.3) builds it up out of hyperelementary levels (it induces an inductive norm on $T_{\mathfrak{A}}$). The basic facts about $T_{\mathfrak{A}}$ are what we should expect.

Lemma 3.3.5. *Let* $\mathfrak{A} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$.

1. $T_{\mathfrak{A}}$ *is inductive on* $\mathfrak{M}$.
2. *For* $\sigma < O(\mathfrak{M})$, $T_{\mathfrak{A},\sigma}$ *is hyperelementary.*
3. (The Weak T Property.) $(T_{\mathfrak{A}}(\ulcorner\varphi\urcorner) \vee T_{\mathfrak{A}}(\ulcorner\neg\varphi\urcorner)) \rightarrow (T_{\mathfrak{A}}(\ulcorner\varphi\urcorner) \leftrightarrow \langle\mathfrak{A}, T_{\mathfrak{A}}\rangle \models \varphi)$.

Note we do not know that $T_{\mathfrak{A}}$ is non-hyperelementary if $\mathfrak{A} \neq \mathfrak{M}$. This will depend on the particular structure $\mathfrak{A}$. For extremely simple structures, like $\langle M\rangle$ (or more properly $\langle M, M^{k_1}, \ldots, M^{k_n}\rangle$), every inductive relation is elementary, and so $T_{\mathfrak{A}}$ will be in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$.

Now, we want to investigate how a structure codes up its own contribution to the truth relation, by examining how a structure $\mathfrak{A}$ articulates $T_{\mathfrak{A}}$.

Theorem 3.3.6. *Let* $\eta = O(\mathfrak{M})$.

1. *A structure* $\mathfrak{A} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ *articulates* $T_{\mathfrak{A},\alpha}$ *in* $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ *for* $\alpha < \eta$.
2. *No structure in* $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ *with* $\|T_{\mathfrak{A}}\| = \eta$ *articulates* $T_{\mathfrak{A}}$ *in* $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. *In particular,* $\mathfrak{M}$ *does not articulate* $T_{\mathfrak{M}}$ *in* $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$.
3. *There are structures in* $\mathbb{H}\mathrm{YP}_{\langle\mathfrak{M},Tr\rangle}$ *that articulate their* $T_{\mathfrak{A},\eta}$ *in* $\mathbb{H}\mathrm{YP}_{\langle\mathfrak{M},Tr\rangle}$.
4. *In particular, any structure* $\mathfrak{A}$ *in* $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ *articulates its* $T_{\mathfrak{A}}$ *in* $\mathbb{H}\mathrm{YP}_{\langle\mathfrak{M},Tr\rangle}$.

One way to prove this is by a lemma which is of some interest in its own right. It tells us that we can map sentences in $T_{\mathfrak{A}}$ into $\mathcal{K}_{\mathbb{H}\mathrm{YP}_{\mathfrak{M}}}$ in a way that respects levels.

Definition 3.3.7. Let $\ulcorner\varphi\urcorner \in T_{\mathfrak{A}}$. The *rank* $|\varphi|$ of φ is the least α such that $\ulcorner\varphi\urcorner \in T_{\mathfrak{A},\alpha}$.

We can map sentences of rank α into sentences of $\mathcal{K}_\alpha$ true in $\mathfrak{A}$.

Lemma 3.3.8. *Let* $\mathfrak{A} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$*. There is a* $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$*-recursive* $1-1$ *map* $\natural\colon T_{\mathfrak{A}} \to \mathcal{K}_{\mathbb{H}\mathrm{YP}_{\mathfrak{M}}}$ *such that for each* φ*,* $\mathfrak{A} \models \varphi^{\natural}$*. Furthermore, if* $|\varphi| = \alpha$ *then* $\varphi^{\natural} \in \mathcal{K}_{\alpha}$*.*

Proof. We define $\natural$ by recursion on $\alpha = |\varphi|$. For $|\varphi| = 0$, $\langle \mathfrak{A}, \emptyset \rangle \models \varphi^*$, which implies that φ contains no occurrences of $\mathsf{T}_{\mathfrak{A}}$. Set $\varphi^{\natural} = \varphi = \varphi^* \in \mathcal{K}_{\omega\omega} = \mathcal{K}_0$. Then clearly $\mathfrak{A} \models \varphi^{\natural}$. Observe also that as $T_{\mathfrak{A},0}$ is hyperelementary, so is its image under $\natural$.

Suppose the result for sentences of rank $\alpha < O(\mathfrak{M})$. Let $S_\alpha = \{\psi^{\natural} \mid |\psi| \leq \alpha\} = \natural[T_{\mathfrak{A},\alpha}]$. We may suppose S_α is a hyperelementary subset of $\mathcal{K}_\alpha$. Thus, $\bigvee\{v \doteq \ulcorner\dot{\psi}\urcorner \mid \psi \in S_\alpha\}$ is a formula of $\mathcal{K}_{\alpha+1}$. For each φ of rank $\alpha+1$, let $\varphi^{\natural}$ be the result of replacing each occurrence of $\mathsf{T}_{\mathfrak{A}}$ in φ^* by $\bigvee\{v \doteq \ulcorner\dot{\psi}\urcorner \mid \psi \in S_\alpha\}$. The result is a sentence of $\mathcal{K}_{\alpha+1}$. As $T_{\mathfrak{A},\alpha+1}$ is hyperelementary by Lemma (3.3.5), so is its image $\natural[T_{\mathfrak{A},\alpha+1}]$.

Induction on formulas shows that each $\varphi^{\natural}$ is true in $\mathfrak{A}$. It is clear from the construction that $\natural$ is $1-1$ and defined by Σ recursion on $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$, making it $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$-recursive. Indeed, the operations involved in defining $\natural$ are syntactic, so with care, it can be defined by Δ recursion. ■

From the lemma, the proof of Theorem (3.3.6) follows.

Proof of Theorem (3.3.6). $\natural$ can be inverted in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ on any hyperelementary subset of its codomain, so Lemmas (3.3.5) and (3.3.8) guarantee that each $\mathfrak{A}$ articulates $T_{\mathfrak{A},\alpha}$.

The rest of the theorem follows easily. ■

Theorem (3.3.6) shows how from the perspective of $\langle \mathfrak{M}, \mathit{Tr} \rangle$, we can think of each structure $\mathfrak{A}$ in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ as implicitly containing the predicate $T_{\mathfrak{A}}$, or (relying on Lemma 3.3.4) its $\mathfrak{A}$-section of Tr. From $\langle \mathfrak{M}, \mathit{Tr} \rangle$, we can thus think of each such $\mathfrak{A}$ as really looking like $\langle \mathfrak{A}, T_{\mathfrak{A}} \rangle$. The information that each structure gives can be combined, to reconstruct an internal version of a truth relation within a context.

Definition 3.3.9. Let $A \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ be a set of structures in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ that articulate their $T_{\mathfrak{A}}$ in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. Let the *internal A-truth relation* of $\mathfrak{M}$, $\widetilde{Tr}^{A}_{\mathfrak{M}} = \coprod_{\mathfrak{A} \in A} T_{\mathfrak{A}} = \{\langle \pi(\mathfrak{A}), \ulcorner\varphi\urcorner \rangle \mid \mathfrak{A} \in A \wedge \ulcorner\varphi\urcorner \in T_{\mathfrak{A}}\}$.

Any context $\mathfrak{M}$ has a maximal internal truth relation. But of particular interest to us is the case where the context is $\langle \mathfrak{M}, \mathit{Tr} \rangle$ and $A = \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$: the internal $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$-truth relation of $\langle \mathfrak{M}, \mathit{Tr} \rangle$. Call this relation $\widetilde{Tr}^{\mathfrak{M}}_{\langle \mathfrak{M}, Tr \rangle}$ (the internal version of the truth relation for $\mathfrak{M}$ from the perspective of $\langle \mathfrak{M}, \mathit{Tr} \rangle$). From Theorem (3.3.6) we have:

Corollary 3.3.10. $\widetilde{Tr}^{\mathfrak{M}}_{\langle\mathfrak{M},Tr\rangle}$ *is well-defined, and is in* $\mathbb{H}\mathrm{YP}_{\langle\mathfrak{M},Tr\rangle}$*, but* not *in* $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$.

Remark 3.3.11.

1. It is clear from Theorem (2.5.14) that:

$$\widetilde{Tr}^{\mathfrak{M}}_{\langle\mathfrak{M},Tr\rangle} \subsetneq \coprod_{\mathfrak{A}\in\mathbb{H}\mathrm{YP}_{\langle\mathfrak{M},Tr\rangle}} T_{\mathfrak{A}}.$$

   We may embed an internal truth relation $\widetilde{Tr}$ into the corresponding genuine Tr by using the second translation of the proof of Lemma (3.3.4), which replaces each occurrence of $\mathsf{T}_{\mathfrak{A}}(x)$ with $\mathsf{Tr}(\pi(\dot{\mathfrak{A}}), x)$. Under this embedding, we have:

$$\widetilde{Tr}^{\mathfrak{M}}_{\langle\mathfrak{M},Tr\rangle} \subsetneq Tr^{\langle\mathfrak{M},Tr\rangle}.$$

2. Technically speaking, $\widetilde{Tr}^{\mathfrak{M}}_{\langle\mathfrak{M},Tr\rangle}$ is not the same as $Tr^{\mathfrak{M}}$. However, the difference is incidental. For any element of $Tr^{\mathfrak{M}}$, we may identify a set of elements of $\mathfrak{A}$-sections which witnesses its truth in $\widetilde{Tr}^{\mathfrak{M}}_{\langle\mathfrak{M},Tr\rangle}$. Observe that cross-world comparisons of truth reduce to world-specific ones. For instance, suppose we have $\langle\pi(\mathfrak{A}), \ulcorner\mathsf{Tr}(\pi(\mathfrak{B}), \ulcorner\varphi\urcorner)\urcorner\rangle \in Tr^{\mathfrak{M}}$. This holds because $\langle\pi(\mathfrak{B}), \ulcorner\varphi\urcorner\rangle \in Tr^{\mathfrak{M}}$. Similarly, quantification into the structure position of Tr reduces to a set of instances. The proof of Lemma (2.4.4) shows how to extend this procedure by recursion on $O(\mathfrak{M})$. This can be carried out entirely within $\mathbb{H}\mathrm{YP}_{\langle\mathfrak{M},Tr\rangle}$. Hence, at least from the perspective of $\langle\mathfrak{M}, Tr\rangle$, we may safely ignore the differences between $\widetilde{Tr}^{\mathfrak{A}}_{\langle\mathfrak{M},Tr\rangle}$ and $Tr^{\mathfrak{M}}$.

Remark 3.3.12. In Section (2.6) I defended the use of infinitary languages as a way to invoke definability theory which is faithful to what we can expect of speakers' expressive resources in a context. In particular, I noted that the infinitary language $\mathcal{K}_{\mathbb{H}\mathrm{YP}_{\mathfrak{M}}}$ is no stronger than $\mathcal{K}^+$. The proof of Theorem (2.6.1) showed this by constructing a map from $\mathcal{K}_{\mathbb{H}\mathrm{YP}_{\mathfrak{M}}}$ to $\mathcal{K}^+$.

Lemma (3.3.8) provides a kind of converse to this. For each $\mathfrak{A} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$, $\natural\colon T_{\mathfrak{A}} \to \mathcal{K}_{\mathbb{H}\mathrm{YP}_{\mathfrak{M}}}$. Together, these induces a map on $\widetilde{Tr}^{\mathfrak{M}}_{\langle\mathfrak{M},Tr\rangle}$. The map is truth preserving, in that for $\langle\pi(\mathfrak{A}), \ulcorner\varphi\urcorner\rangle \in \widetilde{Tr}^{\mathfrak{M}}_{\langle\mathfrak{M},Tr\rangle}$, $\mathfrak{A} \models \varphi^{\natural}$.

Relying on the identification of Remark (3.3.11.2), we may take this to show that the expressive resources of $\mathcal{K}^+$ in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ are no stronger than those of $\mathcal{K}_{\mathbb{H}\mathrm{YP}_{\mathfrak{M}}}$. (At least, this can be seen to be so from the expanded context $\langle \mathfrak{M}, Tr \rangle$.) Maps directly defined on $\mathcal{K}^+$ can be generated from the maps Υ of Lemma (3.3.4). For each $\mathfrak{A} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$, $\natural \circ \Upsilon \colon \mathcal{K}^+ \to \mathcal{K}_{\mathbb{H}\mathrm{YP}_{\mathfrak{M}}}$. Each such map is likewise truth preserving on the corresponding $\mathfrak{A}$-section of Tr.

Together with Theorem (2.6.1), these mappings give fairly precise form to the idea that the right infinitary language and the language with semantic relations have the same expressive power.

$\widetilde{Tr}^{\mathfrak{M}}_{\langle \mathfrak{M}, Tr \rangle}$ provides the needed reconstruction from context $\langle \mathfrak{M}, Tr \rangle$ of the semantic relation of the prior context $\mathfrak{M}$. Its construction goes along the lines I sketched informally in Section (3.2). Putting aside incidental syntactic matters (relying on Lemma (3.3.4) and Remark (3.3.11)), the construction starts with each structure from $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$, and finds an articulation of its $\mathfrak{A}$-section. This identifies the semantic contribution of each structure, as implicitly given by the structure. Indeed, from $\mathbb{H}\mathrm{YP}_{\langle \mathfrak{M}, Tr \rangle}$, we can see each structure $\mathfrak{A}$ in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ as looking essentially like $\langle \mathfrak{A}, \mathfrak{A} \upharpoonright Tr \rangle$. The semantic contributions provided by each structure in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ are combined to build $\widetilde{Tr}^{\mathfrak{M}}_{\langle \mathfrak{M}, Tr \rangle}$ as a reconstruction of the semantic relation $Tr^{\mathfrak{M}}$ of context $\mathfrak{M}$. Speakers in context $\langle \mathfrak{M}, Tr \rangle$ can use this to identify the contents of claims made in the prior context $\mathfrak{M}$. This is just what we need to get a better understanding of the Liar inference, and the conclusion l at (B).

3.4 The Paradox Revisited

Now that we have the internal truth relation $\widetilde{Tr}^{\mathfrak{M}}_{\langle \mathfrak{M}, Tr \rangle}$, we can go back and describe what happens in the Liar inference. But first, there is one more loose end that needs to be tie up. In the discussion of Part (1), I used the relation *Exp*, while in Part (2) I switched to the parameterized version *Tr*. As I explained there, the difference is primarily technical, and we can think of *Tr* as parameterizing *Exp*. But we should make sure that the formal developments I have offered really to relate back to the original formulation. This will be important for finishing our assessment of the Liar inference as presented in Section (1.1.4), as it will allow us to apply the formal apparatus of Parts (2) and (3) to it.

We can recover the idea of expressing a proposition, by noting that relative to a fixed context $\mathfrak{M}$, a sentence expresses a proposition just in case it divides the structures in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ into two classes: those in which the sentence holds, and those in which it does not hold. Insofar as we want Tr to express the semantic relation of the context, we thus want to say that φ expresses a proposition just in case $\forall x(Str(x) \to (Tr(x, \ulcorner \varphi \urcorner) \vee Tr(x, \ulcorner \neg\varphi \urcorner)))$. Let us introduce some abbreviations. Say φ is determinate at w if it has a truth value there according to Tr, i.e. $D(w, \ulcorner\varphi\urcorner) \leftrightarrow (Tr(w, \ulcorner\varphi\urcorner) \vee Tr(w, \ulcorner\neg\varphi\urcorner))$. φ expresses a proposition if it is determinate in all worlds: $E(\ulcorner\varphi\urcorner) \leftrightarrow \forall x(Str(x) \rightarrow D(x, \ulcorner\varphi\urcorner))$.

When directly quantifying over propositions, we had as a Liar Sentence the sentence $l = \neg\exists p(\mathsf{Exp}(\ulcorner l\urcorner, p) \wedge \mathsf{Tr}(p))$. We may put this in parameterized form, using $\dot{\mathsf{E}}$ and the *binary* Tr. By the usual diagonal methods, we may obtain a fixed point of the predicate $\dot{\mathsf{E}}(v) \rightarrow \neg\mathsf{Tr}(\pi(\dot{\mathfrak{M}}), v)$: a sentence λ such that $\langle\mathfrak{M}, Tr\rangle \models \lambda \leftrightarrow (\dot{\mathsf{E}}(\ulcorner\lambda\urcorner) \rightarrow \neg\mathsf{Tr}(\pi(\dot{\mathfrak{M}}), \ulcorner\lambda\urcorner))$. Note that $\dot{\mathsf{E}}$ contains a suppressed quantifier over truth conditions (structures), giving us the quantifier made explicit as $\exists p$ in l.

Now, let us return to the Liar inference of Section (1.1.4). First, a calculation following the first part of the Liar inference verifies that in our initial context $\mathfrak{M}$, λ cannot express a proposition, i.e. $\langle\mathfrak{M}, Tr\rangle \models \neg\dot{\mathsf{E}}(\ulcorner\lambda\urcorner)$.

The theory predicts that there is now a context shift to a new context $\langle\mathfrak{M}, Tr\rangle$, with an expanded domain of truth conditions given by $\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}$. Relative to this new context, one can argue that the Liar sentence λ is true. After all, we have $\neg\dot{\mathsf{E}}(\ulcorner\lambda\urcorner)$, so we have $\neg(\dot{\mathsf{E}}(\ulcorner\lambda\urcorner) \wedge \mathsf{Tr}(\pi(\dot{\mathfrak{M}}), \ulcorner\lambda\urcorner))$, i.e. $\dot{\mathsf{E}}(\ulcorner\lambda\urcorner) \rightarrow \neg\mathsf{Tr}(\pi(\dot{\mathfrak{M}}), \ulcorner\lambda\urcorner)$. But this just λ.

This conclusion is made in the expanded context $\langle\mathfrak{M}, Tr\rangle$. But crucially, it draws on the conclusion $\neg\dot{\mathsf{E}}(\ulcorner\lambda\urcorner)$ made in the *previous context* $\mathfrak{M}$. For this reasoning to be valid, it must keep the interpretation of the semantic expressions $\dot{\mathsf{E}}$ and Tr fixed as the context shifts. To do this, speakers need to be able to make sense from $\langle\mathfrak{M}, Tr\rangle$ of the semantic relations $\dot{\mathsf{E}}$ and Tr as they were interpreted in the prior context $\mathfrak{M}$. They may do so within $\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}$ by constructing the internal truth relation $\widetilde{Tr}^{\mathfrak{M}}_{\langle\mathfrak{M}, Tr\rangle}$. By Remark (3.3.11), we may treat this as an internal reconstruction of $Tr^{\mathfrak{M}}$. Let us simplify the notation and write $\widetilde{Tr}$ for the internal reconstruction of $Tr^{\mathfrak{M}}$. We also have an internal $\widetilde{E}(\ulcorner\varphi\urcorner) \leftrightarrow \forall\mathfrak{A} \in \mathbb{HYP}_{\mathfrak{M}}(\widetilde{Tr}(\pi(\mathfrak{A}), \ulcorner\varphi\urcorner) \vee \widetilde{Tr}(\pi(\mathfrak{A}), \ulcorner\neg\varphi\urcorner))$. Using these, we can construct an *internal interpretation* of λ as it was interpreted in $\mathfrak{M}$. Write this as $\tilde{\lambda}$.

It should be stressed, this is not to suggest that there are distinct *sentences* λ and $\tilde{\lambda}$. The same syntactic item Tr appears in both. Rather, it is only to mark how the same sentence is being interpreted across shifting contexts. (Speakers could introduce defined symbols for the internal semantic relations, but that is not what is under interpretation in the Liar inference.) In the expanded context $\langle\mathfrak{M}, Tr\rangle$, to preserve the interpretation of λ as it was used in $\mathfrak{M}$, it must be interpreted as $\tilde{\lambda}$.

Now, the correct observation to be made in context $\langle\mathfrak{M}, Tr\rangle$ is that in context $\mathfrak{M}$, we had $\neg\dot{\mathsf{E}}(\ulcorner\lambda\urcorner)$. In the expanded context, this says $\neg\widetilde{E}(\ulcorner\lambda\urcorner)$. It follows that $\neg(\widetilde{E}(\ulcorner\lambda\urcorner) \wedge \widetilde{Tr}(\pi(\dot{\mathfrak{M}}), \ulcorner\lambda\urcorner))$. Hence, $\tilde{\lambda}$.

This conclusion may be correct because $\widetilde{Tr}$ is no longer the semantic relation of the expanded context $\langle\mathfrak{M}, Tr\rangle$. As Corollary (3.3.10) makes clear, it is an ordinary definable relation. Both $\widetilde{Tr}$ and $\widetilde{E}$ are relations on M that are elements of $\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}$. Working in $\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}$, we may simply observe that $\ulcorner\lambda\urcorner \notin \widetilde{E}$. We may reach this conclusion by following the paradox, or we may directly inspect the construction of the $\mathfrak{A}$-sections of Tr. This reveals that, for instance, that we cannot have $\langle\pi(\mathfrak{M}), \ulcorner\lambda\urcorner\rangle$ or $\langle\pi(\mathfrak{M}), \ulcorner\neg\lambda\urcorner\rangle$ in the $\mathfrak{M}$-section of Tr, and hence (relying on Lemma (3.3.4)), they cannot be in $\widetilde{Tr}$. We thus witness $\neg\widetilde{E}(\ulcorner\lambda\urcorner)$. It follow by logic $\neg(\widetilde{E}(\ulcorner\lambda\urcorner) \wedge \widetilde{Tr}(\pi(\dot{\mathfrak{M}}), \ulcorner\lambda\urcorner))$.

As elements of $\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}$, $\widetilde{Tr}$ and $\widetilde{E}$ are relations on M definable by formulas of $\mathcal{K}_{\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}}$. (This is implicit in the proof of Theorem (2.5.5). Since Tr is inductive on $\mathfrak{M}$, it is definable in $\mathcal{K}_{\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}}$, which allows us to then define other elements of $\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}$. Cf. [1, VI.5.10].) Hence, if we like, we can conclude $\mathfrak{M} \models \tilde{\lambda}$. Likewise, elements of $\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}$ are fully definable by formulas of $\mathcal{K}^{+}$ over $\langle\mathfrak{M}, Tr^{\langle\mathfrak{M}, Tr\rangle}\rangle$ (that is, both them and their complements are definable), by (2.4.6) and a theorem of McGee [38, 5.21]. Hence, if we think of the internal semantic relations as defined this way, we can conclude $\langle\mathfrak{M}, Tr^{\langle\mathfrak{M}, Tr\rangle}\rangle \models \tilde{\lambda}$, where, importantly, $Tr^{\langle\mathfrak{M}, Tr\rangle}$ is the interpretation of Tr in context $\langle\mathfrak{M}, Tr\rangle$. This is not paradoxical, as I said, because $\widetilde{Tr}$ is not the truth relation of this context. Hence, we no longer have the fixed-point property for $\widetilde{Tr}$. We have the fixed-point property for the real truth relation of this context, $Tr^{\langle\mathfrak{M}, Tr\rangle}$, but we have already seen in Remark (3.3.11.1) that this is a distinct relation from $\widetilde{Tr}$.

Because the internal semantic relations are in $\mathbb{HYP}_{\langle\mathfrak{M}, Tr\rangle}$, and so available to speakers in $\langle\mathfrak{M}, Tr\rangle$ while not being the genuine semantic properties of $\langle\mathfrak{M}, Tr\rangle$, the Liar sentence can express a true proposition in the expanded context $\langle\mathfrak{M}, Tr\rangle$. What proposition is this truth? The apparatus of internal truth predicates allows us a transparent statement. The true proposition is the one expressed by interpreting the Liar sentence as $\tilde{\lambda}$ in context $\langle\mathfrak{M}, Tr\rangle$. This says that that the Liar sentence did not express a true proposition in $\mathfrak{M}$. This is just what we should have for reasoning across contexts. We start with a claim in context $\mathfrak{M}$, and hold its interpretation constant across the context shift to draw a conclusion from it. This is just like what we did in the 'whale' example of Section (3.2).

What truth conditions go with this claim? First, let us look at this informally. Sentences that make semantic claims, like the Liar sentence or the sentence $\dot{\mathsf{E}}(\ulcorner\lambda\urcorner)$, make claims about a range of worlds, or about a specific world. As such, their truth values will generally be independent of the world of evaluation. With this in mind, let us look again at the Liar inference. First, $\neg\dot{\mathsf{E}}(\ulcorner\lambda\urcorner)$ is asserted in $\mathfrak{M}$ (corresponding to context (A) from Sections (1.1.3) repeated in Section (3.2)). We have seen this is true relative to $\mathfrak{M}$. Hence, it is true in all worlds available in $\mathfrak{M}$. It thus expresses the proposition whose truth conditions are $\{\mathfrak{A} \mid \mathfrak{A} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}\}$. In context $\mathfrak{M}$, the truth conditions of $\neg\dot{\mathsf{E}}(\ulcorner\lambda\urcorner)$ are the set of all available worlds.

This tells us what is expressed by λ in $\langle\mathfrak{M}, Tr\rangle$ (corresponding to context (B)), i.e. the truth conditions of $\tilde{\lambda}$. The conclusion $\tilde{\lambda}$ follows from the conclusion expressed by $\neg\dot{\mathsf{E}}(\ulcorner\lambda\urcorner)$ in context $\mathfrak{M}$, and so will be true in just the same worlds. Hence, in $\langle\mathfrak{M}, Tr\rangle$, λ expresses the proposition whose truth conditions are $\{\mathfrak{A} \mid \mathfrak{A} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}\}$. This is what we should expect from $\tilde{\lambda}$. It says that λ did not express a true proposition according to the semantics of context $\mathfrak{M}$. It is $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$, and the structures in it, that tell us about the semantics of $\mathfrak{M}$. The claim is thus true in all the worlds available there.

This analysis of the truth conditions of $\tilde{\lambda}$ could be presented more formally. We have already seen that $\ulcorner\lambda\urcorner \notin \widetilde{E}$, and hence, with the definability resources of $\mathbb{H}\mathrm{YP}_{\langle\mathfrak{M}, Tr\rangle}$, $\mathfrak{M} \models \tilde{\lambda}$. We could follow out the algorithm for cross-world comparisons suggested in Remark (3.3.11.2) to conclude that $\tilde{\lambda}$ is true in each $\mathfrak{A} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$, more or less tracking the informal analysis just presented. But the apparatus of internal truth relations offers us a way to make the conclusion much more vivid. I noted above that the internal truth relation is build by looking at the contribution of each structure individually, and that from the perspective of $\langle\mathfrak{M}, Tr\rangle$, we can see each structure in $\mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ as looking like $\langle\mathfrak{A}, \mathfrak{A} \uparrow Tr\rangle$ for purposes of internal interpretation. With this in mind, we can observe that for no $\mathfrak{A} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$ do we have $\langle\pi(\mathfrak{A}), \ulcorner\lambda\urcorner\rangle \in \mathfrak{A} \uparrow Tr$ or $\langle\pi(\mathfrak{A}), \ulcorner\neg\lambda\urcorner\rangle \in \mathfrak{A} \uparrow Tr$. Hence, our internal interpretation shows us that for each $\mathfrak{A}$, $\langle\mathfrak{A}, \mathfrak{A} \uparrow Tr\rangle \models \neg\dot{\mathsf{E}}(\ulcorner\lambda\urcorner)$. This implies that under the internal interpretation, $\langle\mathfrak{A}, \mathfrak{A} \uparrow Tr\rangle \models \lambda$ for each $\mathfrak{A} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}$. The internal interpretation thus shows us that $\tilde{\lambda}$ holds in each world available in the initial context $\mathfrak{M}$, just as our informal analysis suggested.

It worth remarking upon the sense in which $\tilde{\lambda}$ is a new proposition. After all, its truth conditions are $\{\mathfrak{A} \mid \mathfrak{A} \in \mathbb{H}\mathrm{YP}_{\mathfrak{M}}\}$. This is coextensive with the necessary proposition of context $\mathfrak{M}$. But observe, it is not the necessary proposition of the expanded $\langle\mathfrak{M}, Tr\rangle$, and indeed, relative to this context, we can see it as expressing information about the previous context. This proposition, as such, is not available in the previous context. Hence, the paradox is resolved much as I proposed in Part (1).

3.5 The Nature of the Hierarchy

We can now see better the role hierarchy plays in the theory I have proposed. As I mentioned, context, in the guise of extraordinary context dependence, does the heavy lifting. It is context dependence that makes a difference between utterances (A) and (B). It is context that makes it the case that there are more truth conditions available at (B), and so there is something for the Liar sentence to express there. On the other hand, as I discussed in Section (3.1), this already implicates a kind of emergent hierarchy. This hierarchy gets put to use when we come to explain what proposition λ expresses in the new (B) context $\langle \mathfrak{M}, \mathit{Tr} \rangle$. For to figure this out, we have to consider reasoning across a context shift. To do this, we have to reconstruct the semantics of the (A) context $\mathfrak{M}$ in the expanded context. This leads us to build the internal truth relation $\widetilde{Tr}$, to use to interpret λ. Using it, we can correctly come to the conclusion that the true assertion in $\langle \mathfrak{M}, \mathit{Tr} \rangle$ has the content that in the (A) context $\mathfrak{M}$, λ does not express a true proposition. This has truth conditions the set of worlds available in the (A) context.

The emergent hierarchy is essential to enabling us to do this. It provides the extra worlds needed for there to be a proposition for λ to express in the expanded $\langle \mathfrak{M}, \mathit{Tr} \rangle$. But furthermore, we require the extra strength of $\mathbb{H}\mathrm{YP}_{\langle \mathfrak{M}, \mathit{Tr} \rangle}$ to develop the internal semantic relations. The same reasons there are more truth conditions at all in the expanded context $\langle \mathfrak{M}, \mathit{Tr} \rangle$ enables us to build the internal interpretation of λ.

I shall devote the rest of this section to some comments on the nature of the emergent hierarchy and the role it plays in my proposed resolution of the paradox.

In some ways, the account of $\tilde{\lambda}$ I have offered is quite traditional. We wind up, in context $\langle \mathfrak{M}, \mathit{Tr} \rangle$, with two truth relations at work: the genuine semantics of the context $\mathit{Tr}^{\langle \mathfrak{M}, \mathit{Tr} \rangle}$, and the internal reconstruction of the semantics of the prior context $\widetilde{Tr}$. The non-paradoxical conclusion is one that makes reference to the lower-level truth relation: the internal reconstruction of the semantics of the prior context.

On the other hand, I do want to stress that we arrive at this by non-traditional means. We do not start by positing a hierarchy of truth relations, as some hierarchical approaches do. Nor do we posit an index on the truth relation (an approach I rejected in Section (1.1)). Rather, we arrive at the hierarchy by meeting demands placed upon us by context dependence, and exploiting resources made available by extraordinary context dependence. It is only the demands of reasoning across context that lead us to interpret the Liar sentence in any other way than via the semantic relations of the current context. It is the expanded expressive resources of the expanded context which make it possible to formulate the internal truth relation, and it is the expanded domain of truth conditions of the context which provides a proposition for the Liar sentence to express. Hence, I maintain, the hierarchy arises naturally, given the rather drastic pressures put upon context by paradoxical reasoning.

As it is driven by context dependence, the hierarchy I propose does not posit a lexical ambiguity of the truth relation. Though ultimately an internal truth relation is used in interpretation, this no more indicates a lexical ambiguity of the truth relation than the 'whale' example of Section (3.2) indicates a lexical ambiguity of the demonstrative 'that'. The meaning of the truth relation remains constant, just as the meaning of 'that' remains constant. But in both cases, when looking across context shifts, we have to be careful to reconstruct the right context-dependent value.

Outside of reasoning across context shifts, the expression Tr should have the interpretation assigned to it by the current context. In the context $\langle \mathfrak{M}, Tr \rangle$, this is simply $Tr^{\langle \mathfrak{M}, Tr \rangle}$. We depart from this, by appealing to an internal truth relation, only because of the demands of reasoning across context. In this situation, the internal semantic relation $\widetilde{Tr}$ is no longer the genuine semantic relation of $\langle \mathfrak{M}, Tr \rangle$. If it were, we would simply have the paradox back. But as with any hierarchy, this opens up the possibility of further steps. In the expanded context $\langle \mathfrak{M}, Tr \rangle$, we have a truth relation $Tr^{\langle \mathfrak{M}, Tr \rangle}$, which is the real semantic relation of the context. With it, we can simply repeat the process of generating a step in the hierarchy, leading to a new context, a newly expanded domain of truth conditions, and new internal truth relations. The hierarchy is open-ended.

It should be also stressed that the hierarchy which emerges, in all its aspects, is one which allows a great deal of self-application of semantic relations within levels. This is clear from the development of Part (2), which used Kripkean techniques to ensure a great amount of self-application.

When we look at the basic idea of a hierarchy generated by context dependence, it may well appear that we should expect incomparable contexts, and so not have a proper, linearly ordered hierarchy. It is a striking consequence of the machinery developed here that it predicts a linear ordering of domains of truth conditions. This follows from Theorem (2.5.1), which tell us that we get the same effect on $\mathbb{HYP}_{\mathfrak{M}}$ by adding any inductive non-hyperelementary relation to $\mathfrak{M}$. I believe this is the right result. What leads to the expansion of a domain of truth conditions, as I see it, is brute expansion in the complexity of resources context provides to speakers. It does not matter precisely which more complex expression is made salient.

It may well be possible to find mathematics to provide non-linear orderings of domains. Perhaps one could make use of incomparable hyperdegrees over the natural numbers to do so. But I do not see any natural interpretation of the combinatorial subtlety involved. Rather, it appears to me that the mathematics I have made use of here correctly models the phenomena, and linearity should be accepted as a consequence.

To conclude this section, it will be useful to look back once more to the situation-based proposal of Barwise and Etchemendy [2] I discussed in Section (1.1.5). As I said before, there are a number of similarities between their proposal and mine. As we have already seen in Section (1.2.2), salience structures are structured entities that bear important similarities to situations in their role as contexts. Moreover, in Section (3.3), I observed that from the expanded context $\langle \mathfrak{M}, Tr \rangle$, we could see each world of the previous context as looking like $\langle \mathfrak{A}, \mathfrak{A} \upharpoonright Tr \rangle$. This looks very much like the situation-based idea of expanding a situation by adding a semantic property to it, leading to an expanded domain of truth conditions relative to which an appropriate Liar sentence can be true.

In Section (1.1.5) and (2.7), I considered and rejected the situation-based approach as a starting point. In starting with situations and then tacking on semantic relations, it seemed to me too much like simply assuming a Tarskian division into levels to begin with. As I have stressing, a hierarchical structure emerges in my view, and what emerges does have these features in common with the situation-based approach. But rather than simply postulating that we can non-trivially append semantic relations to some situations, I hope I have explained how this sort of structure emerges from the workings of language. I did not assume any situation-theoretic apparatus: truth conditions are taken to be structures that reasonably model possible worlds. But once we see that this does not preclude extraordinary context dependence, we can see that we still can have domains of truth conditions expanding as context shifts. This means speakers can in fact have increased expressive resources in some contexts. In the face of reasoning across a context shift in the Strengthened Liar, they are faced with a task of reconstructing the semantics of the previous context. Along the way, they need to identify the contributions of individual worlds in the previous context to the semantics, and this is what allows them to think of those worlds as having situation-like structure in the expanded context. Thus, speakers wind up *discerning* semantic structure in some truth conditions. They do this by way of the additional expressive resources the expanded context provides. In contrast to the situation-based theory, the additional structure discerned is in the truth conditions originally available. It is not the basis for the expansion of the domain of truth conditions; rather, it is an aspect of how speakers in the expanded domain may make sense of the prior domain. But nonetheless, it is fair to say that what speakers are enabled to discern in the expanded context is situation-like structure. But we need not start with a situation-based theory to understand how and why speakers do this.

The hierarchy I propose is primarily a hierarchy of contexts, which induces a hierarchy of domains of truth conditions, which in turn induces a hierarchy of semantic relations. Reasoning across the shifts in context the hierarchy provides leads to the introduction of internal truth relations in some contexts, and hence to a more traditional hierarchical feature of multiple interpretations of the truth relation, as well as to some similarities with situation-based theories.[23]

3.6 Reflection and Hierarchies

Many hierarchical proposals have taken their cue from the idea that the second conclusion of the Liar inference is somehow made by way of *reflecting* upon a previous conclusion or a previous context. Stepping into a position of reflection is stepping up a level in a hierarchy. I basically agree. In other work [17], I have argued that this sort of phenomenon is widespread, and does not lead to an objectionable fragmentation of a concept. I thus have tried to argue that there mere fact that we encounter a hierarchy of this sort is not an objection *per se*. I shall conclude by addressing how the specific proposal I have advanced here relates to the general idea of reflection.[24]

It is clear that in some respects, the hierarchical proposal I have offered captures some kinds of reflection. The conclusion at (B) is made by way of an internal truth relation, which reconstructs the semantics of the prior context (A). This is a kind of reflection on the semantic potentials of sentences in context (A). The proposition expressed at (B) bears this out, as it winds

[23] A number of criticisms of hierarchical approaches assume a much more traditional hierarchy than the one I have developed. This is clear for the classical objections of Kripke [32], as I have build into the hierarchy as much self-application of the truth relation as Kripke himself provides. I also believe it holds for some more recent criticisms, such as those of Simmons [48, 49], though a more thorough discussion of this will have to wait for another occasion.

[24] The theme of reflection appears in a number of works, including the classic papers of Kripke [32] and Parsons [43]. It also appears in more recent work, such as Simmons [48, 49]. As I already mentioned, Simmons' work embraces some aspects of hierarchy, while rejecting others. I have argued in [17] that reflection is the core of the hierarchy. For this reason, as well as those mentioned in Footnote (23), I emphasize the hierarchical aspects more than Simmons does; though as I mentioned in that note, a more thorough discussion will have to wait for another occasion.

up 'saying' that the Liar sentence cannot express a true proposition in the (A) context.

But this fails to address some aspects of the Liar inference as natural-language phenomena, as we have been considering it here. We simply do not seem to see any sort of explicit reflection on the status of a previous context. Speakers following the Liar inference do not explicitly step back and reflect in this manner. Even if speakers can do so, it is not evident why this process should be at work in the Liar inference. As I have been stressing, we can see the crucial step in the Liar inference, from (A) to (B), as an instance of linguistic context shift; albeit extraordinary, but still a natural-language phenomenon. We do not generally expect context shifts to require such explicit reflection on the part of speakers. Why should this extraordinary context shift be any different?

What is happening, I believe, is that the phenomenon of extraordinary context dependence has the effect of making such reflection 'automatic', and not transparent to speakers. The basis for this effect, as I have been stressing in this part, is the role of reasoning across contexts. Speakers are lead to adjust for a shift in context, as they often do. This can be transparent, but it can also be a significant achievement. In this case, the achievement required is the construction of an internal truth relation. This is indeed an instance of reflective thinking about a previous context, in that it is a reconstruction of the semantics of that context. That speakers can sometimes succeed in this without fully realizing it is not all that surprising. Linguistic competence, and the pragmatic ability to use language, are complex abilities, the details of which are often not transparent to speakers. That speakers can sometimes, roughly, recognize the 'reflective' or simply 'odd' character of the Strengthened Liar, shows that they may get a more or less accurate glimmer of what sort of natural-language processes are happening underneath the surface of their linguistic achievements.

If my picture is right, introducing semantic notions like truth or expression into a discourse can induce context shifts. These context shifts can have rather drastic consequences. If the semantic notion is complex enough, the shift in context can expand the domain of truth conditions from which propositions may be constructed. This induces a hierarchy, which has several aspects. When we consider inferences such as the Liar, we can run into paradox if we do not take due note of the hierarchy induced.

I hope to have shown that both the phenomenon of extraordinary context shift, and the induced hierarchy, are not *ad hoc* or unmotivated. There is good reason to see context as shifting by adding salient items. In some cases, there is good reason to see this as expanding the domain of propositions.

And there is good reason to see reasoning across this shift as relying on internal reconstructions of prior semantic relations. This is the core of the contextual-hierarchical approach I propose.

References

[1] J. Barwise. *Admissible Sets and Structures.* Springer-Verlag, Berlin, 1975.

[2] J. Barwise and J. Etchemendy. *The Liar.* Oxford University Press, Oxford, 1987.

[3] T. Burge. Semantical paradox. *Journal of Philosophy*, 76:169–198, 1979. Reprinted in [37].

[4] J. P. Burgess. The truth is never simple. *Journal of Symbolic Logic*, 51:663–681, 1986.

[5] D. Büring. Topic. In P. Bosch and R. van der Sandt, editors, *Focus: Linguistic, Cognitive, and Computational Perspectives*, pages 142–165. Cambridge University Press, Cambridge, 1999.

[6] R. Carnap. *Meaning and Necessity.* University of Chicago Press, Chicago, 1947.

[7] N. Cutland. Model theory on admissible sets. *Annals of Mathematical Logic*, 5:257–289, 1973.

[8] S. Feferman. Some applications of the notions of forcing and generic sets. *Fundamenta Mathematicae*, 56:325–345, 1965.

[9] S. Feferman. Toward useful type-free theories, I. *Journal of Symbolic Logic*, 49:75–111, 1984. Reprinted in [37].

[10] H. Gaifman. Operational pointer semantics: Solution to self-referential puzzles I. In M. Y. Vardi, editor, *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 43–59. Morgan Kaufmann, Los Altos, 1988.

[11] H. Gaifman. Pointers to truth. *Journal of Philosophy*, 89:223–261, 1992.

[12] H. Gaifman. Pointers to propositions. In A. Chapuis and A. Gupta, editors, *Circularity, Definition and Truth*, pages 79–121. Indian Council of Philosophical Research, New Delhi, 2000.

[13] P. C. Gilmore. The consistency of partial set theory without extensionality. In T. J. Jech, editor, *Axiomatic Set Theory*, volume 13, Part II of *Proceedings of Symposia in Pure Mathematics*, pages 147–153. American Mathematical Society, Providence, 1974.

[14] M. Glanzberg. The Liar in context. *Philosophical Studies*, 103:217–251, 2001.

[15] M. Glanzberg. Supervenience and infinitary logic. *Nous*, 35:419–439, 2001.

[16] M. Glanzberg. Topic and discourse. *Mind and Language*, 17:333–375, 2002.

[17] M. Glanzberg. Truth, reflection, and hierarchies. *Synthese*, forthcoming.

[18] B. J. Grosz. The representation and use of focus in a system for understanding dialogs. In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, pages 67–76. Morgan Kaufmann, Los Altos, 1977. Reprinted in [21].

[19] B. J. Grosz, A. K. Joshi, and S. Weinstein. Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics*, 21:203–225, 1995.

[20] B. J. Grosz and C. L. Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12:175–204, 1986.

[21] B. J. Grosz, K. Sparck Jones, and B. L. Webber, editors. *Readings in Natural Language Processing*. Morgan Kaufmann, Los Altos, 1986.

[22] A. Gupta. Truth and paradox. *Journal of Philosophical Logic*, 11:1–60, 1982. Reprinted in [37].

[23] I. Heim. *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. dissertation, University of Massachusetts at Amherst, 1982. Published by Garland, New York, 1989.

[24] P. G. Hinman. *Recursion-Theoretic Hierarchies*. Springer-Verlag, Berlin, 1978.

[25] R. C. Jeffrey. *The Logic of Decision.* McGraw-Hill, New York, 1965.

[26] H. Kamp. A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, and M. Stokhof, editors, *Truth, Interpretation, and Information*, pages 1–41. Foris, Dordrecht, 1984.

[27] L. Karttunen. Discourse referents. In J. D. McCawley, editor, *Notes from the Linguistic Underground*, volume 7 of *Syntax and Semantics*, pages 363–385. Academic Press, New York, 1976.

[28] H. J. Keisler. *Model Theory for Infinitary Logic.* North-Holland, Amsterdam, 1971.

[29] H. J. Keisler. Forcing and the omitting types theorem. In M. D. Morely, editor, *Studies in Model Theory*, pages 96–133. Mathematical Association of America, Buffalo, 1973.

[30] R. C. Koons. *Paradoxes of Belief and Strategic Rationality.* Cambridge University Press, Cambridge, 1992.

[31] A. Kratzer. An investigation of the lumps of thought. *Linguistics and Philosophy*, 12:607–653, 1989.

[32] S. Kripke. Outline of a theory of truth. *Journal of Philosophy*, 72:690–716, 1975. Reprinted in [37].

[33] R. K. Larson. Implicit arguments in situation semantics. *Linguistics and Philosophy*, 11:169–201, 1988.

[34] D. Lewis. Scorekeeping in a language game. *Journal of Philosophical Logic*, 8:339–359, 1979. Reprinted in [35].

[35] D. Lewis. *Philosophical Papers*, volume 1. Oxford University Press, Oxford, 1983.

[36] P. Ludlow. The adicity of 'believes' and the hidden indexical theory. *Analysis*, 56:97–102, 1996.

[37] R. L. Martin, editor. *Recent Essays on Truth and the Liar Paradox.* Oxford University Press, Oxford, 1984.

[38] V. McGee. *Truth, Vagueness, and Paradox.* Hackett, Indianapolis, 1991.

[39] L. McNally. On recent formal analyses of topic. In J. Ginzburg, Z. Khasidashvili, C. Vogel, J.-J. Lévy, and E. Vallduví, editors, *The Tbilisi Symposium on Logic, Language and Computation*, pages 147–160. CSLI, Stanford, 1998.

[40] Y. N. Moschovakis. *Elementary Induction on Abstract Structures.* North-Holland, Amsterdam, 1974.

[41] M. Nadel. More Löwenheim-Skolem results for admissible sets. *Israel Journal of Mathematics*, 18:53–64, 1974.

[42] M. Nadel. Scott sentences and admissible sets. *Annals of Mathematical Logic*, 7:267–294, 1974.

[43] C. Parsons. The Liar paradox. *Journal of Philosophical Logic*, 3:381–412, 1974. Reprinted in [37].

[44] P. Portner and K. Yabushita. The semantics and pragmatics of topic phrases. *Linguistics and Philosophy*, 21:117–157, 1998.

[45] T. Reinhart. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica*, 27:53–94, 1981.

[46] C. Roberts. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Ohio State University Working Papers in Linguistics*, 49:91–136, 1996.

[47] G. E. Sacks. *Higher Recursion Theory.* Springer-Verlag, Berlin, 1990.

[48] K. Simmons. *Universality and the Liar.* Cambridge University Press, Cambridge, 1993.

[49] K. Simmons. Three paradoxes: Circles and singularities. In A. Chapuis and A. Gupta, editors, *Circularity, Definition and Truth*, pages 333–354. Indian Council of Philosophical Research, New Delhi, 2000.

[50] S. Soames. Direct reference, propositional attitudes, and semantic content. *Philosophical Topics*, 15:47–87, 1987.

[51] R. C. Stalnaker. Assertion. In P. Cole, editor, *Pragmatics*, volume 9 of *Syntax and Semantics*, pages 315–322. Academic Press, New York, 1978. Reprinted in [53].

[52] R. C. Stalnaker. *Inquiry.* MIT Press, Cambridge, 1984.

[53] R. C. Stalnaker. *Context and Content*. Oxford University Press, Oxford, 1999.

[54] J. Stanley. Context and logical form. *Linguistics and Philosophy*, 23:391–434, 2000.

[55] J. Stanley and Z. Gendler Szabó. On quantifier domain restriction. *Mind and Language*, 15:219–261, 2000.

[56] A. Tarski. Der Wahrheitsbegriff in den formalizierten Sprachen. *Studia Philosophica*, 1:261–405, 1935. References are to the translation by J. H. Woodger as "The concept of truth in formalized languages" [57].

[57] A. Tarski. *Logic, Semantics, Metamathematics*. Hackett, Indianapolis, second edition, 1983. Edited by J. Corcoran with translations by J. H. Woodger.

[58] E. Vallduví. *The Informational Component*. Ph.D. dissertation, University of Pennsylvania, 1990. Published by Garland, New York, 1992.

[59] J. van Eijck and H. Kamp. Representing discourse in context. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic and Language*, pages 179–237. Elsevier, Amsterdam, 1997.

[60] B. C. van Fraassen. Presupposition, implication, and self-reference. *Journal of Philosophy*, 65:136–152, 1968.

[61] J. van Kuppevelt. Discourse structure, topicality and questioning. *Journal of Linguistics*, 31:109–147, 1995.

[62] K. von Fintel. *Restrictions on Quantifier Domains*. Ph.D. dissertation, University of Massachusetts at Amherst, 1994.

[63] M. Walker. Centering, anaphora resolution, and discourse structure. In M. Walker, A. Joshi, and E. Prince, editors, *Centering in Discourse*, pages 401–435. Oxford University Press, Oxford, 1997.

[64] G. L. Ward. *The Semantics and Pragmatics of Preposing*. Ph.D. dissertation, University of Pennsylvania, 1985. Published by Garland, New York, 1988.