Cross-Contextual Semantics

Ephraim Glick
University of St Andrews
April 24, 2017

Abstract: An utterance sometimes takes long enough to produce that it extends across semantically relevant contextual changes. Those can include changes in the features of context that determine the extensions of indexicals. The standard semantic approach to indexicality does not accommodate this fact in a natural way. In this paper I sketch an alternative approach that is more appealing.

1 Introduction

On a familiar view about context and communication,[1] the context in which an utterance takes place helps determine the interpretation of the utterance, but the utterance in turn affects the context. As assertions are made and accepted, for example, the context evolves through the asserted propositions becoming presuppositions of the participants. And manifest facts about the speech act itself can also become presupposed, as when the audience registers that a particular person has spoken, that particular words were uttered, and so on. It would be strange if these changes in the common ground occurred only in the spaces between sentential utterances — as soon as someone has uttered a word, the audience can take that fact for granted. Why wait until the end of the utterance? It is not surprising that a number of authors have claimed that context can change mid-utterance.

> Now it is well-known that the context changes as the utterance takes place, to a large extent as a result of the utterance itself. Thus the earlier part of an utterance often affects the context in which the later part is interpreted. This phenomenon of intrasentential variability has been noted many times in the literature on context-change, in connection with conjunctions, conditionals, and the defeasibility of presuppositions (see e.g. Stalnaker 1974, 1978). (Recanati 1996: 453–454)

[1] See Stalnaker (1999) for the classic development of the view.

Stalnaker's statements of the idea also appear in more recent work. He says, for instance, that "the context — the common ground — is changing throughout the process of interpretation" (2014: 79), and this can be used to explain facts about, e.g., the presuppositions of conjunctions. When someone asserts *A and B*, the impact of *A* on the context can make *B* felicitous even if *B* would not have been felicitous on its own. "Even if it remains an open question, midway through the speaker's statement, whether the addressee will accept her assertion that *A*, it will immediately become common ground that the speaker accepts that *A*" (2014: 97). In short, "speakers keep track, in real time, of the evolving information, not only about the attitudes they share, but also about the issues on which they differ" (2014: 98).

The idea of mid-utterance changes in context need not be motivated by a Stalnakerian conception of context. Stanley and Williamson, for instance, argue that correct evaluation of sentences containing multiple quantifiers requires "contextually determined variation in the domain *within* a single utterance of a sentence" (1995: 291). And Kaplan notes that since time passes during an utterance of a single sentence, there might be different contexts used to interpret different parts of a sentence. "If we speak slowly enough (or start just before midnight), a repetition of 'today' will refer to a different day. But this is only because the context has changed" (1989a: 586–7).

Kaplan's example is interesting not just because it is independent of any particular conception of contexts, but because it illustrates context changing in respects that are independent of what utterances have been made. The passage of time that occurs while we speak is not a result of the speaking itself. Similarly, while we speak participants may join or exit the scene and the discussion's location may change. Call this phenomenon *natural context shifting*.

Despite the fact that various theorists have endorsed mid-utterance shifts in context, such shifting has never been carefully investigated and accommodated within an appropriate semantic theory. My goal in this paper is to attempt to do that. Starting with a framework based on ideas from Kaplan (1989b) and Lewis (1970; 1980), I will focus on the interpretation of pure indexicals in examples of natural context shifting. These are, to my mind, the simplest examples, but they already require substantial revision to the standard semantic treatment of context-sensitivity. Further work would be required to to apply the ideas developed below to the sort of utterance-driven context shifting that Recanati and Stalnaker have in mind, but I hope the reader will be able to see roughly how it would go.

My central proposal is one that has been made in passing in the literature but has never been developed: "contextual supplementation works at the level of constituents of sentences or utterances, rather than the level of the sentences or utterances themselves" (Soames 1986: 342). After working out this idea in detail, I will briefly explore some possible competitors and consider some potential extensions of the theory to other

context-sensitive expressions. This brings in complications related to variable-binding and so-called monsters. Those complications are not fully resolved here, but overall the approach proposed, which I will call *cross-contextual semantics*, offers a more natural treatment of the core phenomena than would be possible using more familiar semantic theories.

2 Ephemeral contexts

2.1 The basic phenomenon

Imagine an ordinary conversation during which the participants are walking together. At first, they might be a kilometer from their destination, but eventually they arrive. During that discussion time passes. Perhaps even the date changes if the walk takes place around midnight. And of course the roles of speaker and addressee may be exchanged repeatedly. Given a pre-theoretical notion of context as the situation of an utterance, we can say that the context changes during the conversation. Call a context in this pre-theoretical sense a *P-context*.

It is easy to see that the sort of example above can be adapted to illustrate changes in P-context that occur mid-way through an utterance of a single sentence. Just imagine, for instance, that the conversation is one-sided and consists entirely of a single very long utterance. Perhaps the uttered sentence is formed using dozens of sentential connectives. But instead of considering an utterance of a very long sentence, we could equally well consider a very slowly produced utterance of a normal sentence. If someone speaks slowly enough, she could finish speaking a kilometer away from where she began. Even an utterance of a normal sentence produced at a normal rate of speech could begin and end in different locations if the speaker is traveling at a fairly high velocity (in a car or train, say).

Focusing on temporal location instead of spatial makes all this even more obvious. As Kaplan's example illustrates, during a short conversation or even a single utterance, the day of the week, the month, or the year may change. Cutting things more finely, the fact that no utterance is instantaneous would suggest that we always begin speaking at a different time than we finish, so that natural context shifting is absolutely ubiquitous.

2.2 Contexts in semantics

Now, I want to show that examples like those above have consequences for semantics. So what is important for the discussion below is not simply P-context, but context in the sense relevant to semantics — call it *S-context*. And while it may be obvious from the examples that P-context can change mid-utterance, I am not immediately entitled to infer that this corresponds to a change in S-context. P-contexts and S-contexts may not

be the same thing, so to draw out the desired consequences, I should say more about how they are related and about what S-contexts are.

So, what are S-contexts? Channeling Lewis[2]: If we want to know what a context is, figure out what a context does and then find something that does that. By definition S-contexts play a role in semantics, but what more specific roles do they play? We start with the observation that the interpretations and truth-values of utterances depend on features of the P-contexts in which they occur, e.g. the location of the utterance, the time, and the conversational roles of the people involved. 'I am hungry now', when uttered by A at time t, is true iff A is hungry at t, but when uttered by B at t′, it is true iff B is hungry at t′. For a formal semantic system to capture this dependence, it treats the interpretations of expressions as relative to something incorporating the relevant features of a P-context. The interpretations of the indexicals 'I', 'here', and 'now', for instance, depend on the speaker, location, and time, and via compositional rules this dependence is inherited by whole sentences. An S-context, then, is something that incorporates those features.

This still leaves a question about what sort of object an S-context is. Lewis takes them to be centered worlds, while in Kaplan's formal system they are modelled as tuples — "context is a package of whatever parameters are needed to determine the referent, and thus the content, of directly referential expressions of the language" (1989a: 591, cf. 512). I will adopt the latter idea, taking an S-context to be a tuple including at least an agent, a time, and a location to determine the referents of 'I', 'here', and 'now'. But the tuples should also include a world and perhaps many other components, whatever components are needed to deal with the full range of context-sensitive expressions in English.[3]

The important point here is not really the ontology of S-contexts but rather their function. To capture the fact that utterances can have varying interpretations depending on the settings in which they are produced, we take different settings to correspond to different packages of the relevant features of those settings, and then use those packages to provide different interpretations for indexicals and, via compositional rules, for whole sentences. So a core idea behind the use of S-contexts is that variation in interpretation should be dealt with using variation in S-context.

[2] "In order to say what a meaning is, we may first ask what a meaning does, and then find something that does that." (1970: 192)

[3] Lewis (1980) argues that contexts cannot be tuples because there are too many ways in which the language is context-sensitive. I cannot respond to that argument here, but it is connected to various issues (discussed below) concerning monsters, binding, and semantic composition.

2.3 The evidence for semantics

Having made some observations about the job description of S-contexts, we can turn back to our initial examples of natural context shifting. To show that those examples illustrate mid-utterance shifts in S-context, we need to show that the S-contexts relative to which we interpret two parts of a single utterance can be distinct. We will do this by observing that two occurrences of a single indexical expression in an utterance can receive different interpretations. Hence we will see that our examples of changes in P-context are also examples of mid-utterance changes in S-context.

Start with our example of a conversation taking place during a walk. Imagine that one speaker utters a single very long sentence as the participants are en route to a bus stop. The first portion of the sentence could be 'The bus stop is about one kilometer from here and . . . ', while the final portion of the sentence, produced after an extended period of walking, might be '. . . and the stop is just 50 meters from here'. While such a run-on sentence would be considered poor form, it is a sentence of English nevertheless and it could be uttered truly. And it could only be uttered truly if the two occurrences of 'here' picked out distinct locations. Since one role of S-context is to determine the interpretations of 'here' and other indexicals, we see that the two occurrences of 'here' would have to be interpreted relative to two distinct S-contexts. Thus mid-utterance shifts in S-context are possible.

As noted earlier, other examples can be constructed using speakers who are moving quickly rather than speaking for so long. Riding in even an average car, one could truly utter 'The stop is 100 meters from here, but only 50 meters from here'. Or to take a different sort of example: As a speaker approaches a fire, she might say 'Here it is still quite cool, but over here it is warm'. And it should be clear that analogous arguments can easily be produced by relying on features of S-contexts other than spatial location. Take time, for instance. The lifespan of an S-context, even if it is an extended interval rather than a single moment, must be very limited indeed. For within a very short interval, a speaker's temporal location can change enough to make a difference to the interpretation of the indexical 'now' — imagine someone sitting down after uttering the first of two conjuncts: 'I am standing now, but now I am sitting'. Or imagine a speaker demonstrating how to turn on an appliance by turning a knob halfway through her utterance: 'Right now the oven is off, but now it is on'. To return to Kaplan's example, the day of the week can easily change during an utterance, as when a speaker begins an utterance shortly before midnight, saying 'Today is a Thursday, but today is a Friday'. Such an utterance could be true, which would be impossible if every part of the utterance had to be interpreted relative to single S-context.[4]

[4]For some examples similar to the above, see brief discussions by Radulescu (2012) and by Dever,

It is worth emphasizing that examples of these sorts show that it is not possible to avoid positing mid-utterance S-contextual changes by holding that the spatial or temporal location of a context is a region or interval that is as large as whatever contains the entire utterance. If that view were correct, then 'here' (likewise for 'now') would pick out the same (possibly large) region each time it were uttered within a single sentence. But as pointed out, it is possible for the sort of utterance above to take place in such a way that the whole thing were *true*. It follows that within a single utterance, one token of 'here' can pick out a location one kilometer from a point x and a second can pick out a location 50 meters from x.

An alternative reply to the examples would be to say that an S-context provides multiple locations, multiple times, and so on for any other feature we could use to construct our examples, these multiple components being features of the multiple P-contexts present during different portions of the utterance. This will be briefly discussed in Section 5.1, where I point out some unappealing features of the strategy.

To sum up the argument of this section: Whatever S-contexts are, it is constitutive of them that they must be differentiated when they provide factors that differently affect semantic interpretation. In various examples, the S-contexts used to interpret parts of a sentence uttered before and after a change in P-context *do* differently affect semantic interpretation. Therefore those S-contexts must be different, meaning that there are mid-utterance shifts in S-context.

3 The problem for standard semantic theories

The phenomenon of mid-utterance contextual change has no place in classic approaches to semantics such as those of Kaplan and Lewis. Let us start with Kaplan and state the problem informally at first. Kaplan takes expressions to have both an invariant context-independent value, which he calls *character*, and a varying value that results from putting the character together with a context. The latter sort of value he calls *content*. The character of 'I' for instance, is something that yields the speaker of whatever context it is given, while the character of a whole sentence is something that yields a proposition when provided a context. But it is a presupposition of the framework that for the evaluation of a sentence, a single context must be provided to the sentence's character in order to obtain the proposition. There is no provision for somehow supplying multiple contexts to a single sentence character.

Formally, Kaplan models characters as functions from contexts to contents, and

Pickel, and Rabern (2017). We might also want to consider cases in which people finish each other's sentences, which would amount to mid-utterance changes in speaker and addressee. Soames (2010: 100) mentions this phenomenon in connection with Kaplanian semantics, but does not discuss it at any length.

models contents as functions from world-time pairs, *circumstances of evaluation*, to truth-values. We can represent these functions with the following notation. The character of **s** is $\lambda c.\lambda i.[\![\mathbf{s}]\!]^{c,i}$, and the output of that character given context c' as argument is $\lambda i.[\![\mathbf{s}]\!]^{c',i}$. The output of $\lambda i.[\![\mathbf{s}]\!]^{c',i}$ given circumstance of evaluation i' as argument, the extension of **s**, is $[\![\mathbf{s}]\!]^{c',i'}$. So if **s** is a sentence, $[\![\mathbf{s}]\!]^{c',i'}$ is either 1 (True) or 0 (False). And returning to indexicals, for any c and i, $[\![\mathbf{I}]\!]^{c,i}$ is the agent of c, $[\![\mathbf{now}]\!]^{c,i}$ is the time of c, and so on.[5] The lack of dependence on i illustrates that S-context does all the real work of supplying values for indexicals.

Now notice that $\lambda c.\lambda i.[\![\mathbf{s}]\!]^{c,i}$ is a function with a particular domain. It takes a context as argument, not a pair of contexts, a triple of contexts, or a sequence of any other length. We need a single c to plug into that function. The same goes for $\lambda c.\lambda i.[\![\mathbf{I}]\!]^{c,i}$, $\lambda c.\lambda i.[\![\mathbf{now}]\!]^{c,i}$, and so on for every other character. Furthermore, consider the explanation of utterance truth in a Kaplanian system. Each context c is taken to correspond to a special circumstance i_c, which is simply $\langle w_c, t_c \rangle$, the pair of the world and time of the context, and an utterance at c of **s** is true iff $[\![\mathbf{s}]\!]^{c,i_c} = 1$. We have, for example:

(1) An utterance at c of **I am tired** is true iff $[\![\mathbf{I\ am\ tired}]\!]^{c,i_c} = 1$.

As we see on both sides of the biconditional, Kaplan presupposes that utterances must be evaluated at single contexts.

In response to this, one might have the following thought: The truth-value of a sentence should be determined compositionally using the interpretations of components of the sentence. According to Kaplan, individual words have characters and thus contents in context, so why not supply different contexts to the characters of different components of the sentence, calculating the content of the whole sentence from the resulting contents of the components? Perhaps we do not need to calculate the character of the whole sentence and then supply that character with a single context. I will suggest roughly this strategy later (minus the assumptions about the philosophical roles of characters and contents), but it turns out to be a genuine departure from Kaplan's system. For in that system, composition rules pass down c and i parameters to subsentential expressions so that, in absence of anything that shifts those parameters, all the subsentential expressions will be evaluated with respect to the same c and i as the whole sentence. For instance, the rules yield the following:

(2) $[\![\mathbf{I\ am\ tired}]\!]^{c,i} = 1$ iff $[\![\mathbf{tired}]\!]^{c,i}([\![\mathbf{I}]\!]^{c,i}) = 1$.

[5] I will assume in what follows that the extension of 'here' is a location and the extension of 'now' is a time. If these terms instead function as predicate modifiers, the points below could be restated, but the discussion would become unnecessarily complex.

As is well known, Kaplan also held that there are no context-shifting operators (monsters). So every semantic calculation will resemble (2) in that the same context c will be a parameter for the interpretation of every component of the sentence.

The problem should be clear: The Kaplanian theory predicts that two occurrences of the same indexical-type within a single sentence will have the same content and extension. For instance, if an utterance of 'Right now the oven is not on, but now it is' is evaluated using Kaplan's machinery, then the whole uttered sentence is interpreted relative to a single context c and thus all parts of the sentence are interpreted relative to c. So both occurrences of 'now' are interpreted with respect to c, and will each thus contribute t_c, the time of c, to the truth-conditions of the sentence. Given our examples and arguments from above, the standard approach's prediction is incorrect.

Before proceeding, I want to briefly look at the semantic framework of Lewis (1980), which might fairly be considered the most prominent alternative to Kaplan's. A sketch of the theory will allow me to explain why it faces the same problem. Lewis writes:

> To go beyond syntax, a compositional grammar must associate with each expression an entity that I shall call its *semantic value*. [...] These play a twofold role. First, the semantic values of some expressions, the sentences, must enter somehow into determining whether truth-in-English would be achieved if the expression were uttered in a given context. Second, the semantic value of any expression is to be determined by the semantic values of the (immediate) constituents from which it is built, together with the way it is built from them. (1980: 83)

Lewis proceeds to argue that Kaplan's two-stage process of semantic interpretation is "unwarranted and arbitrary" (1980: 97). Where Kaplan assigns to sentences different contents in different contexts, so that we have "variable but simple" values (they are simple in that they take as argument just a context), we could just as well assign "constant but complicated" values that yield a truth-value when given a pair of a context and an *index* (a tuple of whatever features are shiftable by operators of the language). There is no need for the intermediate step of using characters to generate objects that in turn have truth-values relative to a further parameter. (Indeed, even if we had them, the intermediate objects could not be contents in Kaplan's sense, the objects of assertion and belief.)

Despite Lewis' disagreement with Kaplan, there is much in common between the two frameworks. Both relativize truth to two parameters, context and index or circumstance, and both require an adequate semantic theory to be compositional. Most relevant to our purposes, Lewis follows Kaplan in supposing that the semantic value of a whole sentence should be supplied with exactly one context. We might eventually need more than one index-like coordinate (1980: 99), but there is no suggestion that we could somehow

allow more than one context to play a role in the interpretation of a sentence. We are told that "sentence s is true at context c iff s is true at c at the index of the context c" (1980: 88).

4 A solution

4.1 The general proposal

To begin, let us temporarily set aside the idea of trying to use multiple contexts to interpret a single sentence. Let us think about a simpler adaptation of a two-stage semantic system. For this system, I will continue to assume that the truth-value of a sentence is relative to two parameters. But for now I want to be neutral on some of Kaplan's assumptions about the natures of those parameters. Lewis (1980) has persuaded a number of theorists[6] that, as mentioned just above, an index's components must include various components in addition to a world and a time, e.g. a location and a standard of precision. I want to allow for Lewis' view, along with the supposed upshot that when you supply a context to a "character", you get something that cannot be properly considered to be a content in Kaplan's sense, since the resulting value is not the sort of thing that gets believed or asserted. So if we define characters as functions from contexts to contents, then what we have before inputting a context is not really a character at all.

For lack of a better word, I will continue to call the context-independent value of an expression its *character* even though I am not assuming that the character of a sentence yields a proposition given a context. So I will simply call the post-context values of expressions *intensions*,[7] without assuming that intensions are the objects of belief and assertion. I take intensions to yield truth-values at indices, but again, indices may not be Kaplanian circumstances of evaluation.

Now recall that in Kaplan's original system, the absence of context-shifting devices means that every sub-sentential expression gets interpreted relative to the same context as the whole sentence. While composition rules pass the context parameter down to smaller constituents, nothing is done with that parameter until the level of semantically basic expressions, at which point the characters of the basic expressions use the parameter to determine their contents. We could therefore do just as well in generating intensions for sentences (and therefore truth-values at circumstances or indices) if the only characters we ever considered were those of simple expressions, and if we adjusted our composition

[6] See Rabern (2012) and (2013) for particularly helpful discussion.

[7] Borrowing a term from Rabern (2012), we might instead call context-saturated values *compositional values*. I prefer 'intension' since it is neutral about what sorts of values are governed by the compositional rules of the language.

rules to appeal only to intensions. The intension of the whole sentence can be computed from the intensions of the parts without any information about characters of complex expressions.

There is a complication here, however. If a context parameter were not passed down to simple expressions via rules for interpreting the larger expressions they compose, how would it be determined which context were supplied to the simple expressions' characters? In a sense, of course, on the standard theory nothing determines the context relative to which a sentence is interpreted — in using Kaplanian semantics we evaluate sentences at contexts, but we can choose whatever context we like. In such cases, we could similarly simply choose to evaluate the basic components of a sentence at a certain context and use the results to compose an intension for the sentence. If we conceive of a sentence as a phrase structure tree with basic expressions located at the terminal nodes, then this will seem like a straightforward revision. Instead of looking for the value at a context and index of a structure of nodes, we look for the value at an index of a structure of nodes-in-contexts. The context for each terminal node in a structure is just whatever context would, in Kaplan's original framework, be the context relative to which the whole sentence were evaluated.

In another sense, in the standard system we do not always arbitrarily choose a context for interpretation. Rather, since we want to use our semantic machinery to assign truth-values to utterances, we first consider an utterance occurring in a particular P-context and then the definition of utterance truth tells us to use the S-context corresponding to that P-context in the semantics. As above, we say that an utterance at c of **I am tired** is true iff $[\![\textbf{I am tired}]\!]^{c,i_c} = 1$. How is the S-context of an utterance in a particular P-context determined? Well, in evaluating utterances, we can either focus on an imaginary case of an utterance or a real case. If imaginary, we have simply decided what the P-context should be like, and so we suppose that the utterance is produced by a particular agent at a particular place in a world of a certain sort. This gives us the components of an S-context to use in applying the definition of utterance truth. If instead we are considering a real utterance, then we simply appeal to observed facts about the circumstances surrounding the utterance — P-context facts such as who produced the utterance, where and when it was produced, and so on. That information is not provided by the semantics, but must be gathered by interpreters by independent means. To put the point less epistemically or procedurally: the truth-value of a real utterance is determined by the P-context in which the utterance is produced, but the facts about that P-context are not facts to be predicted by a semantic theory. They are instead empirical facts about the setting of the utterance. So, similarly, if we are revising Kaplan's framework in the way I am suggesting, supplying a context directly to the character of each word, we are free to appeal to the same type of fact about the production of an utterance. When someone utters a sentence, they utter the words in that sentence, and facts about the P-contexts in

which those words are uttered are empirical facts that we take as background to semantic processing.

Now, how do we allow for different S-contexts to be supplied to the characters of different words in a single sentence, accommodating mid-utterance contextual changes? The strategy should now be obvious. Since facts about the P-contexts in which individual words are uttered are empirical facts that, while observable by interpreters, need not be predicted by semantic principles, we can simply accept the datum that a single sentence can have one word uttered in one P-context and another word in a different P-context. Those individual P-contexts each have features that get packaged up into an S-context, and those S-contexts can then be used in the semantics to interpret the words. For example, it may simply be an empirical fact that in the course of producing a particular utterance of *Today is a Thursday but today is a Friday* right around midnight, the speaker has uttered the first token of 'today' on December 10, 2015, and has uttered the second on December 11, 2015. Then we can take the S-contexts for the two tokens to differ in their time component (or date component, if there is a separate date component), and this is what will allow us to predict that the utterance can be true.

4.2 Formal implementation

The idea sketched above can be illuminated with a more formal treatment. I will write the character (context-independent value) of an expression-type using double curly brackets around bold, e.g. the character of 'now' is $\{\{\textbf{now}\}\}$. (The reason for the change in notation will become clearer below — we no longer use a c parameter on our standard denotation brackets.) We can suppose that facts about the characters of basic expressions are not determined compositionally, but are basic facts about the language. It might be useful to think of English as having a *lexicon*, in which appear entries such as the following. (I assume for simplicity that predicates have argument slots for times. A full treatment of tense is beyond the scope of this paper.)

(3) a. $\{\{\textbf{now}\}\} = \lambda c.\lambda i.\ t_c$
b. $\{\{\textbf{I}\}\} = \lambda c.\lambda i.\ a_c$
c. $\{\{\textbf{John}\}\} = \lambda c.\lambda i.$ John
d. $\{\{\textbf{off}\}\} = \lambda c.\lambda i.\lambda t.\lambda x.$ x is off at t at i

We can see from these entries that 'John' and 'off' are not context-sensitive items — whatever context the character is given as argument, the resulting value, the intension, is the same. In contrast, the intensions of 'now' and 'I' can differ depending on the context.

We will next want rules to determine how the interpretations of complex expressions are determined compositionally. How do we calculate, for instance, the intension of

'John is standing now'? Well, there is really no such thing as *the* intension of the sentence, or of 'now', for these things have different intensions in different S-contexts. Different utterances of 'now' can pick out different times, so different utterances of 'John is standing now' can accordingly end up true at different indices. The standard thought is then to say that an expression has an intension relative to a context. We can take that to mean that we get an intension by evaluating an expression's character in a context.

In the standard framework, the relativity of intensions to contexts is represented via parameterization of the interpretation brackets. In $\lambda i.[\![\textbf{I am tired}]\!]^{c,i}$, e.g., one indicates that a context c has already been supplied for evaluation of the sentence. In proceeding further with interpretation of the sentence, the context parameter is passed down from the sentence to its constituents via compositional rules. But as we have noted in the sections above, (i) a ban on monsters means that whatever the context relative to which the whole sentence is interpreted, that same context will be the one relative to which the basic constituents of the sentences are interpreted, and (ii) facts about which context an expression occurs in should be taken as background empirical facts, not facts to be predicted in our semantics. So there is really no point in making our semantic machinery pass a context along via composition rules. We might as well assume that we are simply given expressions in contexts, we evaluate those expressions at those contexts, and then we let our compositional semantics work on those. (We can reconsider the ban on monsters later on.)

The idea just suggested faces a slight obstacle of notation. If we want to supply each basic expression with a context and only afterward apply compositional rules, we won't want to use rules that generate equations looking like (2). We need a way to indicate that each word is being supplied a context independently rather than via relativization of a whole sentence to a context. A natural suggestion would be to superscript each word (or each terminal node) with a meta-language name of a context, giving us, for instance, (4-b) instead of (4-a).

(4) a. $[\![\textbf{I am standing now.}]\!]^{c,i}$
b. $[\![\textbf{I}^{c}\ \textbf{am}^{c}\ \textbf{standing}^{c}\ \textbf{now}^{c}\textbf{.}]\!]^{i}$

Instead of the truth-value, relative to context c and index i, of a phrase structure tree, we take (4-b) to be the truth-value relative to an index i of what we might call a *contextualized phrase structure tree* — a structure resulting from replacing each expression at a terminal node with the value of that expression's character at the context indicated by the superscript. To illustrate the composition of intensions and be a bit more explicit, we could give semantic rules of the sort Kaplan gives. However, I will follow the popular and more realistic semantic model of Heim and Kratzer (1998), assuming that our compositional rules are type-driven and that semantic interpretation applies to binary

branching trees.

(5) *Terminal Nodes (TN)*
If α is a terminal node occupied by $\boldsymbol{e}$, then for any c, i, $[\![\alpha^c]\!]^i = \{\{\boldsymbol{e}\}\}(c)(i)$.

(6) *Functional Application (FA)*
If α is a branching node whose daughters are β and γ, then for any i, if $[\![\beta]\!]^i$ is in the domain of $[\![\gamma]\!]^i$, then $[\![\alpha]\!]^i = [\![\gamma]\!]^i([\![\beta]\!]^i)$.

For occurrences of non-context-sensitive expressions, we can simplify our representations by leaving off any context superscripts. We provide the following definition, where $\boldsymbol{e}$ is a word-type and $c*$ is an arbitrary context:

(7) *Abbreviated Notation for Non-Context-Sensitive Vocabulary (AN)*
If $\{\{\boldsymbol{e}\}\}$ is given in the lexicon, then for any i, if for every context c and c', $[\![\boldsymbol{e}^c]\!]^i = [\![\boldsymbol{e}^{c'}]\!]^i$, then we define $[\![\boldsymbol{e}]\!]^i$ to be $\{\{\boldsymbol{e}\}\}(c*)(i)$.

To illustrate how the intensions of occurrences compose, consider our earlier example of a speaker demonstrating the functioning of the knobs on a stove. For simplicity, suppose the stove is b, and that the speaker says 'b is off now, but b is on now'. Let c' be the context of the first occurrence of 'now', and c'' the context of the second. And to further simplify things, continue supposing that predicates such as 'on' and 'off' each have an argument slot for times. Assuming the obvious simple semantics for 'but' and a structure indicated by square brackets, we have the following:

- $[\![[[\textbf{b is } [\textbf{off now}^{c'}]] \, [\textbf{but } [\textbf{b is } [\textbf{on now}^{c''}]]]]\!]^i = 1$ iff
  $[\![[\textbf{b is } [\textbf{off now}^{c'}]]\!]^i = 1$ and $[\![[\textbf{b is } [\textbf{on now}^{c''}]]\!]^i = 1$.
- $[\![\textbf{off now}^{c'}]\!]^i = [\![\textbf{off}]\!]^i([\![\textbf{now}^{c'}]\!]^i)$ *[From FA]*
- $[\![\textbf{now}^{c'}]\!]^i = \{\{\textbf{now}\}\}(c')(i)$ *[From TN]*
- $\{\{\textbf{now}\}\}(c')(i) = [\lambda c \, . \, [\lambda i \, . \, t_c]](c')(i) = t_{c'}$ *[From the lexicon]*
- $[\![\textbf{off}]\!]^i = \{\{\textbf{off}\}\}(c*)(i)$ *[From AN]*
- $\{\{\textbf{off}\}\}(c*)(i) = [\lambda c \, . \, [\lambda i' \, . \, [\lambda t \, . \, [\lambda x \, . \, x \text{ is off at } t \text{ at } i']]]](c*)(i) =$
  $[\lambda t \, . \, [\lambda x \, . \, x \text{ is off at } t \text{ at } i]]$ *[From the lexicon]*
- So, $[\![\textbf{off now}^{c'}]\!]^i = [\lambda x \, . \, [\lambda i \, . \, x \text{ is off at } t_{c'} \text{ at } i]]$
- Similarly, $[\![\textbf{on now}^{c''}]\!]^i = [\lambda x \, . \, [\lambda i \, . \, x \text{ is on at } t_{c''} \text{ at } i]]$

Given the dictionary entries for names, this gives us the result we want:

- $[\![[[\textbf{b is}\ [\textbf{off now}^{c'}]]\ [\textbf{but}\ [\textbf{b is}\ [\textbf{on now}^{c''}]]]]\!]^{i} = 1$ iff
  b is off at $t_{c'}$ at i and b is on at $t_{c''}$ at i.

In this simple sentence, of course, we do not have any index-shifting operators like 'necessarily' or 'it was the case that'. Furthermore, we have ignored dependence on assignment functions and ignored the possibility that we might need extra composition rules such as predicate modification. However, see below for some discussion. The central point is that we have illustrated a semantic system in which different contexts can be used to determine the intensions and extensions of two words within a single sentence.

4.3 Characters for complex expressions

So far, in the cross-contextual semantics set out above the notion of character applies only to basic expressions. However, we could define characters for complex expressions as follows.

(8) *Composition of Characters* (CC)
If $\boldsymbol{\alpha}$ is a branching node whose daughters are $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, and if for some c', c'', and i, $\{\{\boldsymbol{\beta}\}\}(c')(i)$ is in the domain of $\{\{\boldsymbol{\gamma}\}\}(c'')(i)$, then
$\{\{\boldsymbol{\alpha}\}\} = \lambda c.[\lambda i.[\{\{\boldsymbol{\gamma}\}\}(c)(i)](\{\{\boldsymbol{\beta}\}\}(c)(i))]$.

Since the characters of basic expressions are given in the lexicon, (CC) allows us to define the character of a node whose daughters are two terminal nodes, and from there we can continue applying (CC) to define the characters of more complex trees. For example, since $\{\{\textbf{now}\}\} = \lambda c'.[\lambda i'.\ t_{c'}]$ and $\{\{\textbf{off}\}\} = \lambda c''.[\lambda i''.[\lambda t.[\lambda x.\ x \text{ is off at } t \text{ at } i'']]]$, (CC) tells us that $\{\{\textbf{off now}\}\} =$

$\lambda c.[\lambda i.\ [\ [\lambda c''.[\lambda i''.[\lambda t.[\lambda x.\ x \text{ is off at } t \text{ at } i'']]]](c)(i)]\ ([\lambda c'.[\lambda i'.\ t_{c'}]](c)(i))\ ]$

Reducing, this yields:

$\{\{\textbf{off now}\}\} = [\lambda c.[\lambda i.[\lambda t.[\lambda x.\ x \text{ is off at } t \text{ at } i]]]](t_c) = \lambda c.[\lambda i.[\lambda x.\ x \text{ is off at } t_c \text{ at } i]]$.

We can continue in this fashion to define characters of whole sentences by applying principle (CC) to a sentence's constituents. $\{\{\textbf{b is off now}\}\}$, for instance, turns out to be exactly what you would expect, namely $\lambda c.[\lambda i.[b \text{ is off at } t_c \text{ at } i]]$. And now that we have defined complex characters, we can use standard notation when we want to ignore the possibility of mid-utterance context changes.

(9) *Definition of Standard Notation*
For any expression-type α, let $[\![\alpha]\!]^{c,i} =_{def} \{\{\alpha\}\}(c)(i)$.

Thus, for example, $[\![\textbf{b is off now}]\!]^{c^*,i^*} = \{\{\textbf{b is off now}\}\}(c^*)(i^*) = [\lambda c.[\lambda i.[b \text{ is off at } t_c \text{ at } i]]](c^*)(i^*) = 1$ iff b is off at t_{c^*} at i^*. It is worth noting that, despite these new definitions, our approach diverges from the standard system in that the intension of an utterance *may* not be the result of supplying the character of the uttered sentence with a context. More generally, the intension of a complex part of an uttered sentence may not be the result of providing its character with a context as argument, for sometimes there will be relevant differences between various contexts that must be provided as arguments for words within the complex. If an entire utterance takes place within a single context (perhaps by being a written token held up on a sign), however, then for that case the present approach yields the same result as the standard system.

4.4 The definition of truth

In familiar approaches to semantics, context plays a role not just in the interpretation of lexical items, but in the evaluation of truth or falsity for an utterance. As we saw earlier, in a Kaplanian system an utterance of a sentence **s** made in a context c is true iff the content of **s** in c, evaluated at the circumstance corresponding to c, yields truth. On the view proposed above, there is not always a single context that we can use to determine the relevant point of evaluation for an utterance. Is this not a respect in which the standard approach has an advantage?[8] I will argue not.

First, note that for our machinery, there is no difficulty in interpreting a structure in which every terminal node is taken to occur in the same context. In terms of notation, we would have the same context superscript on every basic expression in the sentence. We simply identify a context for interpretation of all the basic expressions, apply that context via our rule (TN) for terminal nodes, and then proceed with the rest of the interpretation. The truth-conditions generated are the same as on the standard approach.

A second observation that goes along with the first: On the Kaplanian approach, in *every* interpretation of a sentence, there is exactly one context relative to which the sentence is being interpreted. There is no context shifting. So on the Kaplanian approach, in *every* interpretation of a sentence, there is exactly one context relative to which the all the basic expressions in the sentence are being interpreted. As just noted, in our system, we can also easily interpret all the basic expressions relative to the same context (by taking all the occurrences of basic expressions to be in the same context). So we can immediately define truth for any utterance that the original approach could

[8] This concern is voiced by Radulescu (2012) as well as Dever, Pickel, and Rabern (2017), who describe the sort of approach proposed above as including "uncontrolled context shifting".

handle, simply by adopting the usual definition and taking the notation in it as an abbreviation of ours.

Is there any new worry about where the relevant S-context comes from in an application of the definition of truth? No, since any particular application of the definition in the original Kaplanian system, just as in the revised system, presupposes that one has been given a single context in which the utterance occurs. Which context that is is just taken as a background fact. "An utterance of a sentence s in a context c is true iff . . . " — nothing in the system itself explain where that c comes from. If the utterance being considered is a real one, c is just the S-context incorporating relevant features of the setting of utterance (the P-context). If the utterance is imaginary, one just supposes that one has an S-context incorporating relevant features of the imagined P-context. In the cross-contextual semantics above, we are entitled to appeal in the same way to background facts, although we will be more honest about the fact that we are idealizing when we say that an utterance of a whole sentence takes place in a single context.

But what about truth and falsity for the more interesting cross-contextual utterances that have motivated our new approach? Is there a straightforward definition of truth for those? Even if there were not, the new system would not be worse off than the old one — it would just be unclear how to apply its added flexibility to present issue. But there is in fact no *problem* here, just a question that any semantic theory has to answer. To begin to see why, let's extend our previous notation to define utterance intension.

(10) *Definition of Utterance Intension*
If $\mathfrak{u}$ is an utterance of a sentence $[\boldsymbol{\alpha}\ [\boldsymbol{\beta}\ \ldots\ \boldsymbol{\gamma}]]$, where those basic constitutents are uttered in contexts c_1, c_2 $\ldots c_n$ respectively, then the intension of $\mathfrak{u}$ = $\lambda i.[\![[\boldsymbol{\alpha}^{c_1}\ [\boldsymbol{\beta}^{c_2}\ \ldots\ \boldsymbol{\gamma}^{c_n}]]]\!]^i$. We write the intension of $\mathfrak{u}$ as: $\lambda i.[\![\mathfrak{u}]\!]^i$.

When the temporally first basic constituent word of an utterance is uttered in c_1 and the last (perhaps after some semantically relevant contextual changes) is uttered in c_n, let's say that the utterance *spans* c_1 to c_n. Now, we can assume that for any utterance, the entirety of it takes place in a single world, so that any contexts spanned by an utterance will have the same world component. (If we are interested in producing a semantic theory that approximates real-life language use, this assumption will not be problematic. And in any case, the standard approach makes the same assumption.) Now if an index were just a world, the definition of utterance truth would be utterly straightforward. We could take it to be this:

(11) If $\mathfrak{u}$ is an utterance that spans contexts c_1 to c_n, then where w is the world of c_1 to c_n, $\mathfrak{u}$ is true iff $[\lambda i.[\![\mathfrak{u}]\!]^i](w) = 1$.

Of course, Kaplan (1989b) takes a circumstance to be a world-time pair rather than simply a world, while a Lewisian index (1980) is a tuple containing as many features of contexts as are shiftable by operators in the language in question. In either case, the circumstance or index used in the definition of truth is supposed to be closely related to the context of utterance. As Lewis (1980: 86) puts it: "Given a context, there is an index having coordinates that match the appropriate features of that context. Call it the index of the context." For instance, if standards of precision are shiftable, then the index of a context will have as a member the standard of precision operative in that context. And since time is shiftable, the index of a context will have a time as a member. For Lewis, that time will just be the time of the context. We need not hold that every feature of an index is set by a corresponding feature of the context, however — as MacFarlane (2014: 48, 59) notes, the definition of truth may involve quantifying over certain members of the index.[9] There is still a question, though, which is how we treat other parts of indices in our definition of truth once we accept mid-utterance context changes.

To see why this is not a special problem for my proposed semantics, we should note that the real question here is a question for everyone: If the time, standard of precision, location, or other semantically relevant feature of the setting of an utterance (the P-context) changes during the course of the utterance, what is the S-context and corresponding index relative to which the truth-value of the utterance is determined? There is no obvious answer built into the standard approach to semantics any more than there is in the semantics proposed above! I want to suggest that the question is empirical, to be resolved by looking at the judgments we are actually inclined to make about utterances of the sort in question. The investigation could be carried out by a Lewisian or by the proponent of Kaplan's original system just as well as someone who has adopted our proposal above. In what follows, I will sketch some preliminary considerations, illustrating the way I envision the investigation proceeding. By showing how to formulate and test various possible answers to the question, we can demonstrate that there is no straightforward objection to our approach on the grounds of complex circumstance or indices.

Let us focus on the time component of an index, as this is relevant to both Kaplanian and Lewisian approaches. Recall an earlier example of an utterance whose correct interpretation seems to require using different contexts to provide values for two tokens

[9] His example is the assignment function. Since assignments are shiftable, a Lewisian takes them to be components of indices, but MacFarlane doubts that a context provides a function from variables to values. Lewis himself says that "the assignment coordinate cannot naturally be included among features of context" (1970: 195), so the context could therefore not set the value of the assignment component of the index. Rather, we should define truth in a context in terms of truth at every assignment. Whether this is correct or not, the more general point is plausible: A definition of truth might involve quantifying over parts of indices rather than relying on features of the context to set all the initial values of indices.

of 'now': 'I am standing now, but I am sitting now', where the speaker changes position after the first conjunct. What is the relevant index for evaluating that utterance for truth or falsity? More specifically, what time would be a member of the index with respect to which the utterance would be evaluated?

To make an initial observation, consider an utterance of a more simple sentence, 'I am sitting'. If the speaker is not sitting at any time during the interval in which the utterance occurs, the utterance is false. A minimal constraint on utterance truth, then, is that the utterance's intension be true at some time during the interval of the utterance. Further, it seems irrelevant whether the speaker is sitting at any time outside the interval of the utterance. So one possible view is the following.

(12) If u is an utterance in world w that spans contexts c_1 to c_n, then u is true iff for some time t in the interval t_{c_1} to t_{c_n}, $[\lambda i.[\![\mathbf{u}]\!]^i](\langle w, t\rangle) = 1$.

The proposal in (12) may be too weak, however. While perhaps a bit unlikely, we can imagine a quick-moving speaker sitting down for only a split second partway through her utterance of 'I am seated', standing up again before finishing speaking. As audience members, I think we would very reluctant to judge her utterance true. An alternative proposal that accommodates this judgement is the following.

(13) If u is an utterance in world w that spans contexts c_1 to c_n, then u is true iff for every time t in the interval t_{c_1} to t_{c_n}, $[\lambda i.[\![\mathbf{u}]\!]^i](\langle w, t\rangle) = 1$.

This straightforward proposal will look familiar to those working in the semantics of tense. Ogihara (2007: 400) acknowledges that "since it takes time to utter any sentence, an utterance time is necessarily a non-instantaneous interval", and he describes the "standard formal semantic account" of sentences in the simple present tense as treating them as true iff they are "true *throughout* the utterance time". (Unfortunately, Ogihara does not discuss the impact of these ideas on theories of contexts and indexicality.)

Though (13) avoids the problem for (12), it might seem to go too far in the opposite direction. Imagine that a speaker utters 'I am seated', but suppose she begins speaking before sitting down. If she is seated by the time she utters 'seated', her audience may well be inclined to count her utterance true. Interestingly, we get a different judgment in the opposite sort of case. If the speaker is seated at first, but quickly stands up during the course of uttering 'I am seated', her utterance seems to be false. The theorist might resist such judgments, holding that both utterances are false and that the difference in judgments in the two cases is due to pragmatic interference of some kind (rather than due to semantics, or post-semantics). I will not explore this possibility here, and I instead suggest a way to semantically accommodate the apparent datum that utterance truth does not require truth at every time in the interval of the utterance. We could take the

series of P-contexts in which an utterance takes place to determine a relevant sub-interval I$*$ such that the utterance's truth requires truth at every time in I$*$. The new candidate definition of utterance truth is then the following:

(14) If u is an utterance in world w that spans contexts c_1 to c_n, then where I$*$ is the relevant sub-interval of the interval t_{c_1} to t_{c_n}, u is true iff for every time t in I$*$, $[\lambda i.[\![\mathbf{u}]\!]^i](\langle w, t, \dots\rangle) = 1$.

Instead of continuing to assess various proposals, let us step back and reflect on the main point. The important thing is that we have seen that we can formulate clear definitions of utterance truth and can use familiar methods to evaluate those proposed definitions. There is no difficulty created by the changes to semantics proposed so far in this paper. And importantly, it is not as if the standard approach to semantics has an easier time choosing between definitions of utterance truth. If a single context includes times at which the speaker is sitting and times at which the speaker is not sitting, how do we evaluate the utterance of 'I am sitting'? Choosing a response would require sorting through proposals exactly analogous to the proposals just canvassed.

5 Comparison with alternative strategies

While the main aim of this paper is to present a problem and propose one natural solution to it, I want to briefly consider three other possible solutions, or at least ideas that might be thought to lead to solutions. The first is to appeal to referential indices while positing richer contexts. The second is to appeal to monsters, but to otherwise stick more closely than we have to Kaplan's original framework. A third is to appeal to dynamic semantics. I won't purport to conclusively refute each of these views, but I will sketch some worries for them to show why I find the proposal above more attractive.

5.1 Referential indices

To begin with indices: The claim would be that contexts contains sequences of times, sequences of locations, and sequences of other features used in interpreting context-sensitive vocabulary. Instead of taking the context to have shifted between two uses of, say, 'now', we would take the two uses to differ in which member of the sequence of times were relevant to the interpretation of the word. To make this system work, we would need a way to associate a given occurrence of 'now' with a member of the sequence. This would be accomplished by adding subscripts to the words and saying, for example, that in an utterance of 'Right now$_1$ I am sitting, but I am standing now$_2$', the first subscripted 'now' picks out the first time in the context's sequence of times, and the second picks out the second.

I will mention a couple of unappealing features of this view. First, there is in principle no limit to the amount of time a single utterance could take to produce, or the number of significant contextual changes that could take place during a single utterance. In order for the indexing strategy to be worth pursuing even in simple cases, it should generalize to deal with more extreme examples of the sort introduced earlier. But once this is noticed, the strategy begins to seem very implausible. If someone with enough determination utters a sentence that takes two days to produce, do we really want to say that there is a single S-context that must be used to interpret that entire utterance, an S-context that contains hundreds of times? Not unless we have to. For we should keep in mind that the whole point of S-contexts is to explain variation in interpretations of expressions. When the interpretation of an uttered word seems to depend on the P-context of the utterance, we capture that by saying that different P-contexts correspond to different S-contexts, and different S-contexts combine with the word's character to yield different values. If we refuse to use differences in S-context to explain variation in interpretation that is tied to changes in P-context, we are going against the spirit of context-sensitive semantics. Indeed, taken to its logical extreme, the strategy of numerical indices could be used to eliminate the need to ever use more than one context. Just let the single super-context contain infinite sequences of agents, times, worlds, and so on, and add subscripts to every context-sensitive item in the language. This would not be enlightening. It would just obscure the relationship between interpretation and the setting of an utterance.

This last point is connected to some remarks of Dever et al. (2017), who imagine a proposal to index 'I'. Among other observations about why that proposal would be a bad idea, they argue at length that adding indices to expressions is really a way of ambiguating the strings we ordinarily see or hear, and this produces several problems. For instance, the resulting explosion in the size of the lexicon raises worries about the learnability of the language. How could ordinary speakers acquire and distinguish between many phonologically identical lexical items and how could they correctly associate each of those items with the right features of the context?[10]

An interpreter who encounters an utterance of 'now_{31}' must associate the word with the 31st time in a sequence of times contained in the S-context of the utterance. How does she do this? Well, since subscripts aren't pronounced or heard, it is hard to see how they could play any role in the answer. But obviously for the indexing proposal to predict the right extensions for each utterance of 'now_n', the time t_n in the S-context must be the time at which 'now_n' is uttered. And facts about when a word has been uttered are observable by an interpreter. It must be those observable facts that an interpreter relies on to pair utterances of 'now_n' with the appropriate times. But now that we've noted that, referential indices begin to look like artificial devices that

[10]For more discussion of points in the vicinity, see Braun (1996) on demonstratives.

obscure the real phenomenon — all the work in this picture is being done by the fact that interpreters observe different utterances of indexicals being produced at different times. The semantic proposal I've made above looks more natural in comparison. We don't need indices on 'now', and we don't need a whole sequence of times in our contexts. Rather, we simply feed the word's character the time at which the word is uttered.

It might be worth emphasizing that on my view there is only a single lexical item 'now'. (Similarly for other indexicals.) The superscripts used in my proposal *do not* indicate that there are distinct lexical items 'now^{c}' and 'now$^{c'}$'. A c superscript is not part of a lexical item, or indeed part of the object language at all. It is more accurately thought of as a part of the metalanguage, like the i superscript on the denotation brackets. Spelling out the metalanguage notation more explicitly, $[\![[\mathbf{I}^{c}\ [\mathbf{saw}^{c'}\ \mathbf{John}^{c''}]]]\!]^{i}$ is the value at index i of a structure whose top node's left branch is a terminal node occupied by the value of 'I' in context c, and whose right branch is a complex node whose left branch is a terminal node occupied by the value of 'saw' in context c′ and whose right branch is a terminal node occupied by the value of 'John' in context c″. The only lexical items are 'I', 'saw', and 'John'.

5.2 Monstrous semantics

The second alternative to cross-contextual semantics that I want to briefly consider is a Kaplanian semantics with the addition of context-shifting. In this new system, complex expression-types are evaluated relative to an S-context, and composition rules pass that context down for use in evaluating smaller constituents. But then somehow along the way the context gets shifted. One way to do this would be to let the rule for interpreting terminal nodes overwrite the context parameter with a different context, whatever context the terminal node occurs in.

(15) *Kaplanian Monstrous Rule for Terminal Nodes (KMTN)*
If α is a terminal node that occurs in context c′ and is occupied by expression β, then $[\![\alpha]\!]^{c,i} = [\![\beta]\!]^{c',i}$.

Alternatively, one could posit monsters in the object-language that did the same work via a monstrous version of functional application. (The latter would be a composition rule that said, roughly: if one daughter of a node is of a character-seeking semantic type, and the node's other daughter has a character in the domain of that type, the value of the node is the result of feeding the monster the character.) But to do that work, such a monster would still somehow need to shift the S-context to make it correspond to the P-context in which the utterance of 'now' (or other item) was produced. That new S-context would not be determined by the pre-shifted S-context, since natural context shifting is a matter of how the setting of the conversation happens to evolve, not a

matter that can be predicted semantically. (The semantic rules of English do not predict whether we are moving, how fast someone is speaking, and so on.) So we would need many monsters, one for each context of an indexical. And now I worry that one would end up duplicating some of the problems of the referential index strategy. First, there is obviously the semantic and lexical bloat. Second, for the interpreter to determine what the value of a monster were, she would have to rely on empirical observations about the setting in which the neighboring indexical were uttered. But then it looks like the most important work is being done by the fact that the indexical is uttered in a particular P-context. I prefer to just let such facts play a more direct role in interpretation.

A more general concern is that semantically shifting the context via monstrous devices looks highly unnatural. It is a datum that over the course of an utterance, a speaker may change her location, objects may enter or exit the scene, and of course each successive word of an utterance is uttered later than the previous. It is odd to insist that in interpreting an utterance, we begin by evaluating the whole uttered sentence with respect to a single context but then use *semantic machinery* to shift the context for us. It shifts on its own. We interpreters just need to keep track of how it shifts. Furthermore, until we encounter the monstrosity taking place in the interpretation of non-terminal nodes, the initial context parameter gets idly passed down only to be overwritten. What is the point?

To emphasize, I am not rejecting the possibility of monsters. I am rejecting the idea that natural context shifting should be explained monstrously. If we do want to add monstrosity into the cross-contextual semantics I have proposed, we can do that. Recall that we defined characters for complex-expression types, so all we would need to do to create a monster is give an expression a character-seeking intension and introduce the monstrous version of functional application. For more on these matters, see 6.2.

5.3 Dynamic semantics

The third alternative approach to consider is dynamic semantics, which has context-change at its core. In dynamic semantic theories, the fundamental notion is not truth-conditions, but instead potential to update the context or update the current discourse representation — for classic discussions, see Heim (1983), Kamp (1984), and Groenendijk and Stokhof (1991). A feature of such systems that might be thought similar to features of the cross-contextual semantics presented above is that it predicts that different portions of an utterance can be interpreted with respect to different contexts. A sentence of the form $\phi \wedge \psi$, for instance, is interpreted by updating the previous context with the semantic value of ϕ, then updating the result with the semantic value of ψ. More complex updates are used for sentences containing other connectives. For instance, in a system like Heim's, interpreting a sentence of the form $\phi \rightarrow \psi$ involves creating

a secondary, temporary context by updating the prior context with ϕ, updating that secondary context with ψ, and then removing from the prior context any assignments that failed to survive the latter update.

Of course dynamic systems have been designed with the main goals of treating phenomena like presupposition and anaphora rather than the interpretation of terms like 'now', 'here', 'today', and so on. But someone might think that a dynamic system would be a more natural starting point than a static system in the attempt to develop tools to deal with the phenomena we began with. Could one be adapted or extended to do all the work that we want a cross-contextual semantics to do? There are reasons to think that this would not be easier than the strategy we took above.

First, in dynamic semantics the order in which sub-sentences are used to update a context will not always correspond to the temporal ordering in which those sub-sentences could be uttered by a speaker. Instead of 'If ϕ, then ψ', consider an utterance of the form 'ψ, if ϕ'. Here ϕ is uttered later than ψ, and this must be taken into account if ϕ and ψ both contain a token of the same context-sensitive element whose interpretation depends on a contextual factor that changes mid-way through the utterance. But the utterance will be interpreted just like anything of the form $\phi \rightarrow \psi$. ψ will end up being used to update a temporary context to which ϕ has previously been added. So the order in which sub-sentences are dealt with by dynamic semantics can differ from the real-time ordering of those sub-sentences. Since it is the latter ordering that is relevant to the phenomena we began with, it seems that dynamic semantics does not offer any natural tools to help us.

Second, in familiar dynamic semantic systems there is no built-in system for interpreting different components of an atomic sentence with respect to different contexts. Operators and connectives (or concatenation of sentences, treated like conjunction) produce the desired effect of interpreting one sub-sentence in a different context than another sub-sentence, and if not for the problem above this might work for certain examples of mid-utterance context shifts. But consider an utterance of 'My property extends from here to here', where the speaker quickly moves between the two occurrences of 'here', letting her location at the time of each tokening determine the location picked out by the token. There are no logical connectives or other expressions that dynamic semantics could use to predict a contextual update between the two occurrences of 'here'. We would need additional machinery to get the effect we want.

One could try to develop a dynamic system that offered word-by-word updates rather than updating only via the semantics of connectives,[11] but there would still be the first problem noted above. Furthermore, the updating rules would have no predictive power with respect to the phenomena we are interested in. When a speaker changes

[11]For a system of this sort, see the work of Maria Bittner, e.g. her (2001).

location over the course of an utterance, or time simply passes, the change is not one that is induced by anything semantic. The context shifts we are interested in are natural changes in the setting of a conversation, and whether or how they occur depends on nature and on the whims of the language users. Our semantic machinery should be able to take advantage of these changes, as the proposed cross-contextual semantics does, but we cannot pretend that the changes occur in some rule-governed way. (Even in the case of time — while perhaps time is always passing, the rates of speech in two different utterances of the same sentence might be dramatically different.)

6 Possible extensions of the theory

So far I have focused only on indexicals, but if there are mid-utterance changes in context, we should expect such changes to have effects on context-sensitive items in general. For this reason, I want to briefly explore the possibility of broadening the scope of the theory. The ideas below are fairly tentative, and should be thought of as providing a starting point for future work rather than as providing complete semantic treatments of the phenomena.

6.1 Recurring demonstratives

Easy as it is to produce examples of utterances containing multiple, differently-valued occurrences of 'here' or 'now', it is even easier to find examples involving demonstratives. Any utterance of 'That is bigger than that', 'that looks like that', or 'that *is not* that', etc., involve different tokens of a single lexical item, 'that', being used to refer to different objects. In principle, there seems to be no limit to the number of such tokens, as a speaker might demonstrate any number of different objects over the course of an utterance. This phenomenon, like the phenomenon we began with involving pure indexicals, poses a problem for the standard approach to semantics. Known as "the problem of recurring demonstratives", it has been tackled with strategies analogous to the monstrous strategy and the indexing strategy discussed above. Since recurring demonstratives obviously bear some resemblance to recurring indexicals, it is worth suggesting a way to apply cross-contextual semantics here as well.

Since I am concerned with the *recurrence* rather than with the semantics of demonstratives per se, I will assume a very simple semantics. Consider the "Bare Bones" theory[12] of demonstratives according to which S-contexts include demonstrata and demonstratives have as their characters functions from S-contexts to those demonstrata. I will leave aside questions about the role of demonstrations except for two assumptions.

[12] See Caplan (2003) and Salmon (2002).

(i) If an object is demonstrated in a P-context, that object is the demonstratum in the corresponding S-context. (ii) If different objects are being demonstrated at t_1 and t_2, then there are distinct P-contexts at t_1 and t_2. For our purposes, we can then consider the following simple lexical entry:

(16) $\{\{\textbf{that}\}\} = \lambda c.[\lambda i.$ the demonstratum in $c]$

Following Braun (1996) and Dever et al. (2017), the problem emerges from the assumption that there is a single unambiguous lexical item, 'that'. For once we have assigned a character to that item, as in (16), a standard semantics will predict that every token of it within a given utterance will have the same referent. This is because the whole uttered sentence must be interpreted with respect to a single S-context, and that context will be passed down for the interpretation of every sub-sentential element.

To avoid the problem using our cross-contextual semantics, we can say that in true utterances of 'That isn't that' (and similar cases), the successive tokens of 'that' are evaluated at different S-contexts. That claim follows via assumptions (i) and (ii) if there are different objects demonstrated by the speaker at the times of utterance of the two tokens. It is not implausible that that extra condition will always hold in the relevant cases. For to interpret the utterance correctly, the audience must figure out when the speaker is referring to two different objects with two tokens of 'that'. If there were no observable change in what were being demonstrated, they would be unable to do that.

So by accommodating context changes within a single utterance, we can let different occurrences of 'that' pick out different objects despite the word having a single character. To go through a specific example, consider an utterance of 'That is bigger than that', where the first occurrence of 'that' is in c and the second occurrence is in c'. If a is the demonstratum in c and b is the demonstratum in c', then given our semantics, the utterance would be true at an index i iff $[\![\textbf{That}^{c}\ \textbf{is bigger than that}^{c'}]\!]^{i} = 1$, which will be the case iff a is bigger than b at i. This is of course just what we want.

One further comment about demonstratives: As we have simply assumed the Bare Bones theory without argument, a fuller treatment of the problem would require delving into some subtleties. For instance, what is the role of speaker intentions? And the semantic clause given above for 'that' ignores questions about semantic or pragmatic differences between 'that' and 'this'. I have set such complications aside for reasons of space, but further investigation may show that they are orthogonal to the problem of recurring demonstratives in any case. As long as the correct semantics for these terms relies on facts about context, whether they be facts about demonstrations, intentions or other propositional attitudes, or something else, we can hope the general form of the solution above will be useful.

6.2 Recurring bindable expressions

One larger complication about the foregoing is that 'this' and 'that' may be bindable. If so, we need to consider how to combine binding with cross-contextual semantics. Regardless, there are certainly other bindable expressions that can occur multiple times in a single uttered sentence, such as 'he' and 'she'. So let us consider what to do about pronouns.

In cases of recurrences of *unbound* pronouns, it does seem appealing to use mid-utterance context shifts to explain the different values of the occurrences. For instance, if a speaker begins by pointing at Alice and saying 'She is tall…', but then points to Bertha and continues, '…but she is not tall', the entire utterance could be true. And obviously this requires the two occurrences of 'she' to pick out different people. We could explain this in the same way we dealt with 'this' and 'that': the difference in demonstrations during the first and second halves of the utterance indicates a difference in P-context, and so a difference in the S-contexts used to interpret 'she'. If we treat 'she' along the same lines as we treated 'that', e.g. as having as a character a function from a context to the most salient female of that context, then we do not need to use referential indices on the pronouns to explain the variation in reference among recurrences within a single utterance.

But can we avoid numerical indices altogether? Not without a new theory of binding. The standard theory says that in producing a bound reading of a pronoun, we shift the assignment function that is used to pair expressions with objects via their referential indices. What can we say? The question is how to combine a treatment of bound pronouns with the suggested cross-contextual approach to unbound pronouns, and the most obvious strategies are the following.

First, the approach I find most appealing: We could say that natural language contains no variables and its semantics has no need for assignment functions at all. This view has been defended in detail by Pauline Jacobson, e.g. in her (1999) and (2014). The view is too complex to survey here, but suffice to say that it has motivations totally independent of natural context shifting. If there are no variables or assignment functions needed for natural language semantics, the standard uses of those tools do not provide any reason to reject the proposals made above.

Second, we could simply follow Kaplan (1989a: 572) and say that pronouns are lexically ambiguous. When they are bound, they are interpreted via a shiftable assignment function. (We could think of assignment functions as components of the index parameter of evaluation.) When they are unbound, pronouns are treated like demonstratives. If we stick with the Bare Bones theory of demonstratives, this means that the character of an unbound pronoun is a function from an S-context to a component of that context. While straightforward from a technical point of view, the use of ambiguity

is not particularly appealing. Perhaps a uniform approach would be preferable, all else being equal. On the other hand, the semantic roles of bound and unbound pronouns look very different, so it perhaps it is not unreasonable to deal with them with separate strategies. I am unsure which thought is more persuasive here.

Third, we could take an assignment function to be a component of an S-context, and say that S-contexts with different assignment functions correspond to P-contexts in which different individuals are salient enough to be identified as the referents of pronouns. (Whatever the standard story is about how a context determines an assignment function, we can adopt the same story.) Then we would take both bound and unbound pronouns to be interpreted via the assignment function, but bound pronouns would have their referents shifted by a lambda-binder in the syntax, as in Heim and Kratzer (1998). One consequence of this approach is that we would need variable-binders to be monsters. I have no objection here to monsters per se, and others have argued persuasively that we should indeed think of variable-binding as monstrous, both in Kaplan's semantics and in more contemporary theories. (See Rabern (2012) and (2013).) But within the cross-contextual semantics I have proposed, I am not sure how to provide a compositional treatment of the lambda-binders.

Now, I have mainly been writing about pronouns. But if 'this' and 'that' are bindable, they will raise questions exactly analogous to those just discussed. So there is some uncertainty about whether, or how, my treatment of indexicals could be extended to other context-sensitive items. One worry at this point might be that indexicals themselves could turn out to have shifty interpretations.[13] This would mean that the questions above would be relevant even to the core application of cross-contextual semantics. Obviously a full resolution of these issues would require substantial further discussion, but related to the third strategy above I want to emphasize a point made above in 5.2. It seems extremely unappealing to deal with our original examples monstrously, as to do so would be to mistakenly treat the contextual shifts as if they were induced by semantic mechanisms rather than occurring naturally as a result of changes in P-context. Even if cross-contextual semantics could not be extended to demonstratives and pronouns, it would be unfortunate to conclude from this that we cannot use it even for our original examples. I imagine that theorists sympathetic to the comments from Recanati, Stalnaker, and Soames in §1 will agree. Monsters do not seem like an appropriate way to capture the intuitive phenomena.

[13] As has been argued by, e.g., Santorio (2012). I find the examples unconvincing, but going into the details of why would require a very long digression about the nature of possible-worlds semantics.

6.3 Logic

Another issue that remains to be explored is the extent to which one could develop a logic for a language with a cross-contextual semantics. Of course, since the latter has standard semantics as a special case in which changes in context are idealized away, a logic like Kaplan's LD will apply to a restricted version of our language. But an exploration of the unrestricted version will have to wait.[14]

For now, it is worth emphasizing that there is a genuine question about how to characterize good reasoning that takes place over time and across other changes in P-context. It is an interesting datum that one may go wrong by uttering 'Today is a Tuesday; Therefore, today is a Tuesday', or 'If today is a Tuesday, then today is a Tuesday'. It is difficult to see how to produce a logic that deals with this phenomenon, but that is not a reason to think the phenomenon does not exist or to ignore it. Our lack of an appropriate logical system is not an objection to cross-contextual semantics[15] but a motivation for further work.[16]

References

Almog, Joseph, John Perry, and Howard Wettstein, ed. 1989. *Themes from Kaplan*. Oxford: Oxford University Press.

Bittner, Maria. 2001. Surface composition as bridging. *Journal of Semantics* 18:127–177.

Braun, David. 1996. Demonstratives and their linguistic meanings. *Noûs* 30:145–173.

Caplan, Ben. 2003. Putting things in contexts. *Philosophical Review* 112:191–214.

Dever, Josh, Bryan Pickel, and Brian Rabern. 2017. Reviving the parameter revolution in semantics. In *The science of meaning: Essays on the metatheory of natural language*, ed. Derek Ball and Brian Rabern. Oxford: Oxford University Press.

Groenendijk, Jeroen, and Martin Stokhof. 1991. Dynamic predicate logic. *Linguistics and Philosophy* 14:39–100.

Heim, Irene. 1983. File change semantics and the familiarity theory of definiteness. In *Meaning, use, and the interpretation of language*, ed. Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, 164–189. Berlin: Walter de Gruyter.

[14] For some relevant work, see Zardini (2014) and Radulescu (2012; 2015).

[15] As suggested by Dever, Pickel, and Rabern's objection to "uncontrolled context shifting". See also Kaplan's comments about recurring expressions (1989a: 590).

[16] Thanks to audiences at St Andrews, MIT, and the Pacific APA.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Jacobson, Pauline. 1999. Towards a variable free semantics. *Linguistics and Philosophy* 22:117–184.

Jacobson, Pauline. 2014. *Compositional semantics: An introduction to the syntax/semantics interface*. Oxford: Oxford University Press.

Kamp, Hans. 1984. A theory of truth and semantic representation. In *Truth, Interpretation, and Information: Selected Papers from the Third Amsterdam Colloquium*, ed. J. Groenendijk, T. Janssen, and M. Stokhof, 1–41. Dorsrecht: Foris.

Kaplan, David. 1989a. Afterthoughts. In Almog et al. (1989), 565–614.

Kaplan, David. 1989b. Demonstratives. In Almog et al. (1989), 481–563.

Lewis, David. 1970. General semantics. *Synthese* 22:18–67. Reprinted in (Lewis 1983). Page citations from the latter.

Lewis, David. 1980. Index, context, and content. In *Philosophy and grammar*, ed. Stig Kanger and Sven Öhman, 79–100. Reidel.

Lewis, David. 1983. *Philosophical papers, Vol. 1*. New York: Oxford University Press.

MacFarlane, John. 2014. *Assessment sensitivity: Relative truth and its applications*. Oxford: Oxford University Press.

Ogihara, Toshiyuki. 2007. Tense and aspect in truth-conditional semantics. *Lingua* 117:392–418.

Rabern, Brian. 2012. Against the identification of assertoric content with compositional value. *Synthese* 189:75–96.

Rabern, Brian. 2013. Monsters in Kaplan's logic of demonstratives. *Philosophical Studies* 164:393–404.

Radulescu, Alexandru. 2012. The logic of indexicals. Doctoral Dissertation, UCLA.

Radulescu, Alexandru. 2015. The logic of indexicals. *Synthese* 192:1839–1860.

Recanati, François. 1996. Domains of discourse. *Linguistics and Philosophy* 19:445–475.

Salmon, Nathan. 2002. Demonstrating and necessity. *Philosophical Review* 497–538.

Santorio, Paolo. 2012. Reference and monstrosity. *Philosophical Review* 121:359–406.

Soames, Scott. 1986. Incomplete definite descriptions. *Notre Dame Journal of Formal Logic* 27:349–375. Reprinted in (Soames 2009). Page references to the latter.

Soames, Scott. 2009. *Philosophical essays vol. 1*. Princeton: Princeton University Press.

Soames, Scott. 2010. *Philosophy of language*. Princeton: Princeton University Press.

Stalnaker, Robert. 1999. *Context and content*. Oxford: Oxford University Press.

Stalnaker, Robert. 2014. *Context*. Oxford: Oxford University Press.

Stanley, Jason, and Timothy Williamson. 1995. Quantifiers and context-dependence. *Analysis* 55:291–295.

Zardini, Elia. 2014. Context and consequence: An intercontextual substructural logic. *Synthese* 191:3473–3500.