Fixing Reference by Maximising Knowledge

Penultimate Draft, forthcoming in *Ergo*

Abstract

This paper explores the idea inspired by Williamson (2007) that the meaning of a name is the object such that assigning it as referent maximises knowledge. After situating this idea in a charity-based tradition of interpretation and making it more precise, I argue that it suffers from serious problems. I then show why these problems raise a challenge for charity-based frameworks more generally.

1 Introduction

I am interested in the following question: by virtue of what is the referent of a linguistic term fixed? The terms I'm concerned with here are proper names of type e.[1] These are names whose semantic referent/value is an object. A semantic referent is one relatum in the three place relation of *semantic reference*: a relation that semanticists find helpful in positing between a name, context and object. Semantic reference is to be distinguished from *speaker reference*: a four place relation between speaker, term, object and audience. Roughly speaking, a speaker referent is an object that a speaker uses a term to get their audience to think about.[2] For the most part, I'll be ignoring the question of how speaker referents are fixed. So the question I am concerned with here is this: by virtue of what is the semantic value of a proper name of type e fixed? This is a *meta-semantic* question. In this paper I help myself without argument to the content of the agent's various propositional attitudes like knowledge. This parallels Lewis's (1983) strategy in *Languages and Language*. There, he takes for granted the beliefs of the speakers/hearers and asks: by virtue of what do the sentences in, say, the English language get to have the meanings they actually have? Therefore, more foundational questions such as what constitutes knowledge or how the knowledge of an individual (as a physical system) is fixed are

[1] It's possible to run analogues of my arguments using other expressions. However, I focus on proper names of type e for simplicity and vividness.

[2] This is roughly based on some remarks made by Bach (2006, p. 520). The distinction between semantic reference and speaker reference has also been discussed, among many others, by Kripke (1977).

largely beyond the scope of this paper. With the content of the agent's propositional attitudes in place, my main goal is to evaluate Knowledge Maximisation (from now on KM) as an answer to the above meta-semantic question. KM is roughly the view according to which the semantic value of a proper name is the object such that assigning it as semantic value maximises the knowledge expressed by the speaker's sentences.[3]

The paper is structured as follows. In §2 I situate the general idea behind KM in a charity-based framework of interpretation and consider why one might opt for it. In §3 I formulate KM more precisely and evaluate individual and communal variants of it. I observe that all variants suffer from serious problems. Still, in §5, I investigate whether any variant can at least satisfactorily explain the phenomenon of semantic change – a phenomenon that, on the face of it, KM seems well positioned to explain. However, I observe that KM does not ultimately offer such an explanation. In §6 I take a step back and comment on a strategy of charity-based interpretation that attempts to fix *both* linguistic content and knowledge simultaneously via a principle of maximising knowledge. I'll say why the problems I raise are not resolved by opting for this holistic strategy. I'll also say why these issues easily extend to principles that appeal to true belief instead of knowledge. I take stock and conclude in §7.

2 Motivation

The general idea behind KM, most famously associated with Davidson, is that meaning is determined by a principle of charity that

> [assigns] truth conditions to alien sentences that make native speakers [mostly] right when plausibly possible, according, of course, to our own view of what is right. (Davidson 1973, p. 324)

To motivate the above principle Davidson says:

> If I am right in attributing [a particular] belief to you, then [given that beliefs essentially form coherent clusters] you must have a pattern of beliefs much like mine. No wonder, then, I can interpret your words correctly only by interpreting so as to put us largely in agreement. (Davidson 1977, p. 245)

[3]Some criticisms of KM like that by McGlynn (2012) are partly concerned with whether, in certain cases, the agent's beliefs constitute knowledge or not. This sort of issue will not be relevant here as I'll simply be stipulating the agent's knowledge in each case.

The above charity principle is not necessarily the method by which the interpreter/hearer ascertains the content of an utterance or belief. According to Davidson, the principle is how sentences or beliefs get to have the content they have in the first place.[4] Whether and how this content is ascertained by hearers are different issues.

The notion of 'mostly right' in 'assigning truth conditions to alien sentences that make native speakers [mostly] right' can be developed in a variety of ways. Here are a few illustrated with an example by Williamson.[5]

> Emanuel sees a stranger, Celia, standing some distance away. Looking at her face, he judges 'She is F, G, H, . . .'; he ascribes a character and life-history in considerable detail. In fact, none of it fits Celia. By pure coincidence, all of it fits someone else, Elsie, whom Emanuel has never seen or heard of. Does the pronoun 'she' as used by Emanuel in this context refer to Celia or to Elsie? Which of them does he use it to express beliefs about? He accepts 'She is standing in front of me,' which is true if 'she' refers to Celia but false if it refers to Elsie. However, he also accepts 'She is F,' 'She is G,' 'She is H, . . . ,' all of which are false if 'she' refers to Celia but true if it refers to Elsie. We may assume that the latter group far outweighs the former. A principle of charity that crudely maximizes true belief or minimizes error therefore favors Elsie over Celia as the referent of the pronoun in that context. But that is a descriptive theory of reference gone mad. (Williamson 2007, pp. 262-3)

Let's look at this case through the lens of the framework I'm interested in where the agent's propositional attitudes are assumed and it is various meta-semantic proposals about semantic reference fixing that are being investigated.[6] In this framework we can easily see that various charity-based principles are inadequate. Take for example a principle that minimises the expression of false beliefs. This principle is problematic. Suppose Celia and Elsie are our only candidate assignments. Since Emanuel has some true beliefs about Celia and some false beliefs about Celia, then given that Emanuel has no beliefs about Elsie, an assignment of Elsie to the term in question will indeed minimise the expression of false beliefs. This is because that

[4]See a survey of some of Davidson's passages by Glüer (2011, pp. 112-52) that are suggestive of this claim.

[5]Williamson uses pronouns but, as far as I can tell, the same points can be made using proper names.

[6]Williamson and Davidson themselves are more interested in a holistic strategy whereby beliefs and linguistic contents are somehow fixed simultaneously in a foundational theory of meaning and belief. I'll be talking about the ramifications of my discussion for this holistic strategy in §6.

assignment makes Emanuel's sentences express fewer false beliefs by making his sentences express *no* beliefs about Celia or Elsie. Moreover, a principle that enjoins minimising the expression of falsehoods will also be problematic assuming that under the Elsie assignment there are fewer expressed falsehoods than under the Celia assignment. And a principle that maximises the expression of truths will be problematic for a similar reason. Indeed, on the assumption of bivalence, minimising the expression of falsehoods and maximising the expression of truths amount to the same thing.

KM on the other hand will do better than any of the above three charity-based principles and so we already see some initial motivation for it. Emanuel has no knowledge about Elsie because he has no beliefs about Elsie. Therefore, given that he has some knowledge about Celia, the Celia assignment will express more knowledge than the Elsie assignment. The astute reader however will notice that a principle that maximises the expression of true beliefs (TBM) *also* does better than the other three charity-based principles mentioned in the previous paragraph. Therefore, one might wonder how to adjudicate here between KM and TBM. Assuming as is orthodox that knowledge and true belief are not the same thing, this is a good question.[7] However, the problems I raise for KM in the next two sections also arise for TBM and therefore, since both KM and TBM are unacceptable, it's not going to matter all that much to say which one is worse.

3 Individual-level KM

3.1 Counting Sentences

Let $\mathcal{A}_S$ be the set of sentences accepted by speaker S and $\mathcal{K}_S$ the set of propositions known by that speaker. I leave the notion of *accepted sentences* mostly as a primitive though I note, as many theorists maintain, that acceptance is the internal analogue of assertion. Therefore, an accepted sentence may, in some cases, not be uttered. There are many candidate assignment functions that each assign semantic values to all terms in all sentences in $\mathcal{A}_S$. KM1 provides a method of selection.

> **KM1**. The selected assignment is the one under which more sentences in $\mathcal{A}_S$ express knowledge in $\mathcal{K}_S$ than any other candidate assignment.

[7] Obviously we could, in principle, construct a case where one assignment does better on TBM and one on KM. The general structure of the case would be such that under one assignment there are lots of expressed true beliefs that aren't knowledge and on the other assignment there are fewer expressed true beliefs that are knowledge. The question would then be which assignment is the intuitively correct one. However, for the reasons stated in the text, I am not going to investigate cases of this sort.

When I say that *a sentence in $\mathcal{A}_S$ expresses knowledge in $\mathcal{K}_S$* I simply mean that once a candidate assignment function has assigned semantic values to the terms in that sentence the resulting proposition is identical to a proposition in $\mathcal{K}_S$. In order to evaluate KM1 it will be helpful to start by building a very simple toy model with a very small number of accepted sentences to make the discussion manageable. The knowledge in each case will be stipulated and, as with the set of accepted sentences, I'll be working with unrealistically small sets of knowledge. Nevertheless, as I'll try to show, the problems that emerge will generalise to more complex cases with more realistic sets of accepted sentences and knowledge.

Take the following case adapted from Kripke (1977, p. 263).

> **Leaf Raking**. One day our speaker meets Jones (from now on X) at the local market. X introduces himself by saying 'Hello, I'm Jones'. X works as a butcher and becomes well acquainted with our speaker. The next day, our speaker sees Smith (from now on Y) from a distance raking leaves and mistakes him for X. Therefore, our speaker naturally comes to accept the sentence 'Jones is raking leaves'.

I take it that the semantic value of 'Jones' in this case is X only. Of course there is a sense in which the speaker is talking about Y but that is a merely pragmatic sense. Strictly speaking, the speaker's utterance of 'Jones is raking leaves' is false if, say, X is asleep, regardless of what Y is doing.[8] Assume $\mathcal{A}_S$ includes the following sentences only.

(1) Jones is Jones
(2) This guy is this guy
(3) Jones is this guy
(4) This guy is Jones
(5) Jones is raking leaves
(6) This guy is raking leaves
(7) Jones was at the market
(8) This guy was at the market

The demonstrative 'this guy' in some of the above sentences is a perceptual demonstrative whose semantic value is Y. The speaker sees Y raking leaves

[8]Kripke (1977) and many commentators also take this view. The semantic value of 'Jones' is X only and the speaker referents are X *and* Y. Given our working definition of speaker reference in §1, both X and Y are such that a speaker uses 'Jones' to get their audience to think about. Recall that the speaker has mistaken Y for X.

and on that basis accepts 'Jones is raking leaves' and 'This guy is raking leaves' because they accept 'Jones is this guy'. Assume $\mathcal{K}_S$ includes the following known propositions only.

(a) *X=X*

(b) *Y=Y*

(c) *Y is raking leaves*

(d) *X was at the market*

Now take the following candidate assignments of semantic values. f and g assign semantic values to all terms in $\mathcal{A}_S$. f and g are identical except for what they assign 'Jones'. Moreover, they assign all other terms the intuitively correct semantic values. KM1 must select f in order to make the intuitively correct prediction with respect to the semantic value of 'Jones'.

f: f('Jones') = **X**, f('this guy') = Y, etc.
g: g('Jones') = **Y**, g('this guy') = Y, etc.

3.1.1 Knowledge Quantity Problem: The Identity Model

Here is a table that brings the above details together.

Table A: Identity Model Under KM1

Sentence	K ex under f	K ex under g
Jones is Jones	X=X	Y=Y
This guy is this guy	Y=Y	Y=Y
Jones is this guy	·	Y=Y
This guy is Jones	·	Y=Y
Jones is raking leaves	·	Y is raking leaves
This guy is raking leaves	Y is raking leaves	Y is raking leaves
Jones was at the market	X was at the market	·
This guy was at the market	·	·
Total Under KM1	**4**	**6**

Table A presents the knowledge expressed under assignments f and g for each of the target sentences. As illustrated, under assignment f the total number of sentences that express knowledge is 4, and under assignment g it is 6. KM1 will select g and therefore, according to it, the semantic value of 'Jones' is Y. But this is not the intuitive/traditional view. Of course, the set of sentences/knowledge is artificially small but expanding those sets in various ways, as I will now show, does not improve KM1's predictions.

Call a simple subject-predicate sentence 'simple sentence' and the corresponding knowledge 'simple knowledge'. As already observed in Table A above, the quantity of simple knowledge about X is the same as the quantity of simple knowledge about Y. There are also two simple sentences for each predicate. We can preserve these two features of the model in Table A but have those quantities be greater or much greater. In such a model sentences that express identity will always make the total number of sentences that express knowledge under assignment g greater than that under f.[9]

3.1.2 Knowledge Quantity Problem: The Curiosity Model

The above identity model, one might think, is quite unlikely to arise in real life cases owing to the assumed symmetry with respect to the quantity of knowledge. Here is another model that does not have this feature. Suppose that our speaker in **Leaf Raking** had been tasked with observing and memorising as much detail as possible about the way X looks, moves, the texture of his clothes, the way the sun reflects off his hair, skin and so on. The speaker had been tasked with this by, say, a painter who for whatever reason is very interested in X as subject matter. However, it is Y that the speaker ends up observing. The speaker also gets to hear Y talk about their life, interests and family in such a way that the speaker becomes much more knowledgable about Y than they were about X. Suppose all this happens within a day. The speaker's set of knowledge therefore grows considerably containing the propositions *Y is F*, *Y is G*, *Y is H*, etc. This expansion in the set of knowledge corresponds to an increase in the number of accepted sentences containing the name 'Jones' like ⌜Jones is F⌝, ⌜Jones is G⌝, ⌜Jones is H⌝, etc. The expansion in the set of knowledge also corresponds to an increase in the following kind of accepted sentences: ⌜This guy is F⌝, ⌜This guy is G⌝, ⌜This guy is H⌝, etc. (Recall that the speaker accepts 'Jones

[9] Here is another way of expanding the set of accepted sentences/knowledge that I'll mention only to set aside. If an agent accepts any sentence, then it might seem like they accept an infinite number of sentences (e.g., by repeated use of conjunction introduction including conjoining a sentence with itself). Thus, it might seem like the number of sentences that express knowledge under f and the number of sentences that express knowledge under g will always be the same. And therefore, one might think, KM1 is already doomed. I have some sympathy with this concern, but here I want draw out the consequences of KM that do not involve infinite quantities. Thus, at least for the sake of argument, I assume that the operative notion of acceptance can be meaningfully applied to only a finite number of sentences.

is this guy'). Here is a table illustrating the situation with the additions emphasised in bold.

Table B: Curiosity Model Under KM1

Sentence	K ex under f	K ex under g
Jones is Jones	X=X	Y=Y
This guy is this guy	Y=Y	Y=Y
Jones is this guy	·	Y=Y
This guy is Jones	·	Y=Y
Jones is raking leaves	·	Y is raking leaves
This guy is raking leaves	Y is raking leaves	Y is raking leaves
Jones was at the market	X was at the market	·
This guy was at the market	·	·
Jones is F	·	**Y is F**
This guy is F	**Y is F**	**Y is F**
Jones is G	·	**Y is G**
This guy is G	**Y is G**	**Y is G**
Jones is H	·	**Y is H**
This guy is H	**Y is H**	**Y is H**
Total Under KM1	**7**	**12**

As illustrated in Table B, the number of sentences that express knowledge under assignment g is still greater than f's. One crucial reason for this is that assignment g scores 2 for a sentence pair like ⌜Jones is F⌝ and ⌜This guy is F⌝. Assignment f however scores only 1 for that same pair. KM1, once again, generates the wrong semantic prediction by selecting assignment g. Variations in curiosity, it seems, should not have a constitutive impact on a proper name's semantic value, especially when the associated knowledge is accumulated within such a short period of time. However, according to KM1 they do.

3.2 Counting Knowledge

Above, it was the number of sentences that express knowledge that were being counted. However, as illustrated in Table A and B, some sentences expressed the same known proposition. For example, four different sentences expressed $Y{=}Y$ and therefore increased g's score by 4. Some might be tempted to use this observation to revise KM1 to KM2.

> **KM2**. The selected assignment is the one under which more knowledge in $\mathcal{K}_S$ is expressed by sentences in $\mathcal{A}_S$ than any other candidate assignment.

Roughly speaking, KM2 counts the knowledge expressed by the accepted sentences, whereas KM1 counts the sentences that express knowledge. Here is how the identity model in Table A looks like if KM2 is adopted instead.

Table C: Identity Model Under KM2

Sentence	K ex under f	K ex under g
Jones is Jones	X=X	Y=Y
This guy is this guy	Y=Y	Y=Y
Jones is this guy	·	Y=Y
This guy is Jones	·	Y=Y
Jones is raking leaves	·	Y is raking leaves
This guy is raking leaves	Y is raking leaves	Y is raking leaves
Jones was at the market	X was at the market	·
This guy was at the market	·	·
Total Under KM2	**4**	**2**

As illustrated in Table C, under assignment f the number of known propositions that are expressed by the target sentences is 4, whereas under assignment g it is 2. According to KM2 then, the semantic value of 'Jones' is in accordance with assignment f and is therefore X. This is the intuitively correct result. Note that in order to arrive at this result, one must assume a coarse-grained view of propositions under which, for example, the propositions expressed by 'Jones is raking leaves' and 'This guy is raking leaves' under assignment g is identical. This assumption, which I shall grant in what follows, appears to be essential if KM2 is to be an improvement over KM1.

KM2 also generates the right prediction for the earlier curiosity model. Assignment f picks up all the knowledge about X and all the knowledge about Y. This is in contrast to assignment g which only picks up the knowledge about Y. Here is how Table B looks like if KM2 is adopted instead.

Table D: Curiosity Model Under KM2

Sentence	K ex under f	K ex under g
Jones is Jones	X=X	Y=Y
This guy is this guy	Y=Y	Y=Y
Jones is this guy	·	Y=Y
This guy is Jones	·	Y=Y
Jones is raking leaves	·	Y is raking leaves
This guy is raking leaves	Y is raking leaves	Y is raking leaves
Jones was at the market	X was at the market	·
This guy was at the market	·	·
Jones is F	·	Y is F
This guy is F	Y is F	Y is F
Jones is G	·	Y is G
This guy is G	Y is G	Y is G
Jones is H	·	Y is H
This guy is H	Y is H	Y is H
Total Under KM2	**7**	**5**

3.2.1 The Semantic Freedom Problem

KM2 however is problematic. Above, there was no way for assignment g to capture the speaker's knowledge about X and therefore g continued to score less than f. This however need not be the case. One can assume there are sentences *not* containing the name 'Jones' accepted on the basis of seeing X. For example, suppose that when the speaker was at the market looking at X he accepted sentences containing the demonstrative 'that guy' (not 'this guy') like 'That guy is at the market'. And later, when the speaker is looking at Y raking leaves, suppose the speaker also accepted sentences like 'That guy [who was at the market] is this guy [who is raking leaves]'. Assignments f and g assign 'that guy' X because, recall, they only differ with respect to what they assign 'Jones'.[10] Here is how Table D for example looks like when the relevant sentences (emphasised in bold) are added to the table.

[10]The problem under consideration can also be generated by supposing there are sentences with a *name* that's not 'Jones' whose semantic value is X under both assignments.

f: f('Jones') = **X**, f('this guy') = Y, f('that guy') = X, etc.
g: g('Jones') = **Y**, g('this guy') = Y, g('that guy') = X, etc.

Table E: Semantic Freedom Problem Under KM2 & KM1

Sentence	K ex under f	K ex under g
Jones is Jones	X=X	Y=Y
This guy is this guy	Y=Y	Y=Y
That guy is that guy	**X=X**	**X=X**
Jones is this guy	·	Y=Y
Jones is that guy	**X=X**	·
This guy is that guy	·	·
This guy is Jones	·	Y=Y
That guy is Jones	**X=X**	·
That guy is this guy	·	·
Jones is raking leaves	·	Y is raking leaves
This guy is raking leaves	Y is raking leaves	Y is raking leaves
That guy is raking leaves	·	·
Jones was at the market	X was at the market	·
This guy was at the market	·	·
That guy was at the market	**X was at the market**	**X was at the market**
Jones is F	·	Y is F
This guy is F	Y is F	Y is F
That guy is F	·	·
Jones is G	·	Y is G
This guy is G	Y is G	Y is G
That guy is G	·	·
Jones is H	·	Y is H
This guy is H	Y is H	Y is H
That guy is H	·	·
Total Under KM2	**7**	**7**
Total Under KM1	**11**	**14**

As illustrated in Table E, under KM2 f and g have the same score and therefore KM2 fails to select either assignment. No knowledge was missed under either assignment. This result was achieved by assuming there are accepted sentences already in place not containing the name 'Jones' that under g expressed the knowledge about X. And if the knowledge about X can be expressed under g by sentences not containing 'Jones', then assignment g is *free* to assign any semantic value to 'Jones' whatsoever without lowering its score under KM2. This is why the semantic freedom problem is so called.

As illustrated in Table E, the additions also do not improve KM1's predictions: g's score is still greater than f's. The above assumption will not remove the problematic discrepancy generated by KM1.

3.2.2 The Reverse Assignment Problem

So far I've been working with assignments f and g which are identical except for what they assign 'Jones'. Other candidate assignments however are of course possible. Take for example assignment h which is identical to f except that it reverses the values for 'Jones' and 'this guy'. That is, h assigns Y to 'Jones' and X to 'this guy'. The problem with assignment h is that it scores the same as f. If these were the only candidate assignments then KM2 will select neither. Here is how the curiosity model in Table D for example looks like with assignment h instead of g.

f: f('Jones') = $\mathbf{X}$, f('this guy') = $\mathbf{Y}$, etc.
h: h('Jones') = $\mathbf{Y}$, h('this guy') = $\mathbf{X}$, etc.

Table F: Reverse Assignment Problem Under KM2 & KM1

Sentence	K ex under f	K ex under h
Jones is Jones	X=X	Y=Y
This guy is this guy	Y=Y	X=X
Jones is this guy	·	·
This guys is Jones	·	·
Jone is raking leaves	·	Y is raking leaves
This guy is raking leaves	Y is raking leaves	·
Jones was at the market	X was at the market	·
This guy was at the market	·	X was at the market
Jones is F	·	Y is F
This guy is F	Y is F	·
Jones is G	·	Y is G
This guy is G	Y is G	·
Jones is H	·	Y is H
This guy is H	Y is H	·
Total Under KM2	**7**	**7**
Total Under KM1	**7**	**7**

Given the speaker's confusion, h and f have the same score. KM2 only 'cares' about the total quantity of knowledge expressed, not how that knowledge is distributed across the accepted sentences. Crucially, if the speaker wasn't confused then under h none of the simple sentences will express any knowledge and h's score will be lower than f's. Similar remarks apply to KM1 as illustrated in Table F: h and f have the same score.

There are other ways of expanding the speaker's set of accepted sentences/knowledge, but none seem very promising at evading the problems discussed above.[11] Therefore, in the next section, I shall consider communal variants of KM.

4 Communal-level KM

Williamson suggests that a communal approach to fixing linguistic content is more appropriate. He says: 'each individual uses words as words of a public language; their meanings are constitutively determined not individually but socially, through the spectrum of linguistic activity across the community as a whole' (Williamson 2007, p. 98). There are some choice points with respect to what sort of communal knowledge maximisation strategy to pursue. In what follows I consider two that seem like natural places to start. Moreover, the main lessons will likely carry over to other communal strategies.

[11] Here are a few worth noting. Call a sentence that includes two or more simple sentences connected with a boolean connective 'a complex sentence' and the corresponding knowledge 'complex knowledge'. Complex sentences/knowledge can be easily generated from the exisiting set of simple sentences/knowledge. Adding complex sentences will not fix KM2. One reason for this is the semantic freedom problem. Under that problem whatever sentence containing 'Jones' that expresses knowledge under f has an analogous sentence involving another name which expresses that knowledge under g too. Therefore, no knowledge will be lost under g. Moreover, adding complex sentences does not seem to be a promising avenue for KM1. As illustrated with curiosity models, one can massively increase complex knowledge that is only about Y (e.g., $Y\ is\ F \wedge Y\ is\ G$). Due to the speaker's confusion this will correspond to quadruples of complex sentences that will each score, under g, four times the quantity under f. Of course sentences like ⌜Jones was at the market $\wedge$ this guy is F⌝, ⌜Jones was at the market $\wedge$ this guy is G⌝, etc., *only* express knowledge under f. The point however is that there is going to be an even bigger class of complex sentences that *only* express knowledge under g: ⌜Jones is $F \wedge$ this guy is F⌝, ⌜Jones is $G \wedge$ this guy is G⌝, ⌜Jones is $F \wedge$ this guy is G⌝, ⌜Jones is $F \wedge$ this guy is H⌝, etc. (A similar class is one containing sentences like ⌜Jones is $F \vee$ that guy is F⌝, ⌜Jones is $F \vee$ that guy is G⌝, etc). We can make it even harder for f to catch up by making it compete with assignment i which is just like g except that it also assigns more subject terms Y. Note also that expanding the set of knowledge to include propositions not about X or Y is not going to make a difference to the results I've surveyed. Recall that assignments f and g are identical with respect to what they assign all terms except 'Jones'. Therefore, the corresponding increase in accepted sentences will increase the score for both assignments, under KM1 or KM2, by the same quantity.

4.1 Single Set-Pair Strategy (SSP)

Consider a strategy that takes as the set of accepted sentences only one very large set: the set of every sentence such that at least one speaker in the relevant linguistic community accepts it. This strategy will also take as the set of knowledge only one very large set: the set of every proposition such that at least one speaker in the relevant linguistic community knows it. Call this set-pair 'Single Set-Pair' or 'SSP'. Our communal knowledge maximisation principle will then select the assignment that maximises knowledge expressed as if this SSP belonged to a single speaker. Therefore, this strategy will not in general evade the problems presented in §3.[12]

4.2 Multiple Set-Pairs Strategy (MSP)

One might then be tempted to try out a strategy that takes as many sets of accepted sentences as there are speakers and as many sets of knowledge as there are knowing speakers. Each set-pair (constituted by a set of accepted sentences and a set of known propositions) belongs to a different speaker and there is a one-to-one function from set-pairs to speakers. For the purpose of evaluating a candidate assignment this communal strategy looks at each speaker's set-pair to see how the candidate assignment does with respect to expressing knowledge *for that speaker*. The strategy will then favour assignments that maximise knowledge expressed across speakers: each speaker will get a score by that assignment and the selected assignment will be the one that generates the highest total score. In what follows I explore whether MSP is a more promising strategy to fixing linguistic content.

4.2.1 MSP1 – Counting Sentences

Take a communal strategy that counts the number of sentences that express knowledge along the lines of KM1. Take also a communal variant of **Leaf Raking** where the relevant linguistic community is comprised of 100 speakers. Suppose only 30 are confused but each has 100 pieces of simple knowledge about Y (the wrong semantic value) that is all accumulated within a day. This way, the total number of simple sentences that express knowledge about Y under f (the correct assignment) for the confused speakers is $30 \times 100 = 3,000$. This is in contrast to 6,000 sentences under g. Recall

[12]For example, under a communal principle that maximises the number of known propositions expressed along the lines of KM2, the semantic freedom problem will continue to be troubling despite how many speakers are confused. On the other hand, suppose we imagine there being enough confused speakers who have lots of varied knowledge about Y. If so, then it will be easy to generate curiosity style problems for a communal principle along the lines of KM1.

that assignment g exploits the speakers' confusion by scoring double for each sentence pair accepted as a result of the knowledge gained about Y. g's score is greater by $3,000$. We can ensure that f will not catch up as follows. Suppose each of the 100 speakers holds 10 pieces of simple knowledge about X. This corresponds to $1,000$ sentences that g will miss out on and therefore f now has a shortage of at least 2,001 to makeup for in order for it to be selected. But we've set things up so that this shortage will not be made-up. Whatever simple knowledge about Y held by each of the unconfused speakers will correspond to sentences that end up increasing the scores for f and g by the same quantity. Of course unconfused speakers will accept complex sentences like 'Jones is not this guy' which express knowledge only under f but this only amounts to 70 sentences. Similar remarks apply to complex sentences accepted by unconfused speakers on the basis of looking at Y like 'Jones is not standing there'. The upshot of this is that, according to MSP1, semantic values will be influenced in a problematic way by a misinformed minority who pick up lots of knowledge about Y in a day.[13]

4.2.2 MSP2 – Counting Knowledge

Now take a communal strategy that counts the number of known propositions expressed along the lines of KM2. Here, the semantic freedom problem will continue to be troubling whether anyone in the community is confused or not. It will also continue to be troubling no matter how much knowledge each speaker has about Y relative to X. This is because f's and g's scores will be equalised for each set-pair. Recall that under that problem we assume that whatever sentence containing 'Jones' that expresses knowledge under f has an analogous sentence involving another name which expresses that knowledge under g too. Therefore, no knowledge will be missed under g.[14]

[13]Note that the reverse assignment problem will equalise the scores for f and h across the confused speakers. What will be break the tie in favour of f is the score for the unconfused speakers. So long as there is at least one unconfused speaker, that will be sufficient to narrowly break the tie between those assignments. However, it's possible that there are no unconfused speakers. Note also there are variants of the MSP1 strategy under which the selected assignment is not the one with the highest total score but the one with the highest mean, median or mode score. However, as the reader can easily workout for themselves, it's easy to set up examples that make trouble for each of these variants. Moreover, none of these variants will make a difference for variants of MSP that count knowledge instead. This is because, as we shall shortly see, the semantic freedom problem ensures that the scores for f and g are equalised for each set-pair.

[14]Moreover, if all speakers are confused, then the reverse assignment problem will continue to be troubling for MSP2 as this will equalise the total score for f and h.

5 KM & Semantic Change

In this section I present another toy model based on a well-known case in the literature on proper names: the case of 'Madagascar'. This case involves semantic reference change. On the face of it, KM seems to have a ready explanation for cases involving semantic change. Intuitively, the object that becomes the new semantic value of the name is the one that ends up playing a suitably important role in the speaker's life and that will naturally correspond to an increased amount of knowledge about the new object relative to the old. An account of fixing semantic values via maximising the expression of knowledge might seem well positioned to get these cases right.

Evans (1973 p. 195-6) reports from Isaac Taylor's (1898) book *Names and their History*, that 'Madagascar' was, a long time ago, only a name for a section of the African mainland (from now on X). Take the following variant of the case.

> **Madagascar**. Upon first hearing 'Madagascar' from some Arab sailors Marco Polo correctly took it to name X and based on some remarks by the Arab sailors he thought *X was invaded*, *X was peaceful* and *X is an island*. A few days after this Marco Polo gets on a ship and discovers an island (from now on Y). For some reason or other Marco Polo thinks that Y is the island those Arab sailors were naming with 'Madagascar'. That is, he mistakes Y for X. Marco Polo then lands on Y, unpacks and sets himself up for permanent stay. This is the early stage. In the late stage Marco Polo is eighty years old, still using the name 'Madagascar' and has gained much knowledge about Y.[15]

To keep predictions consistent with **Leaf Raking**, we will have to say that when first arriving on Y Marco Polo used the name 'Madagascar' with only X as semantic value. There are many ways of modelling the situation at the early stage. The below is one in which Marco Polo has a bit more knowledge about X than Y. Assume that Marco Polo's $\mathcal{A}_{MP}$ includes the following sentences only.

(9) Madagascar is Madagascar

(10) This island is this island

(11) Madagascar is this island

[15] On standard presentations of the case 'Madagascar' was just a name for a section of mainland Africa – X. After a period of misuse initiated by Marco Polo, the semantic value of the name changed to the island we today call 'Madagascar' – Y.

(12) This island is Madagascar

(13) Madagascar is an island

(14) This island is an island

(15) Madagascar was invaded

(16) This island was invaded

(17) Madagascar was peaceful

(18) This island was peaceful

Suppose also that Marco Polo's $\mathcal{K}_{MP}$ includes the following known propositions only.

(e) *X=X*

(f) *Y=Y*

(g) *Y is an island*

(h) *X was invaded*

(i) *X was peaceful*

Now take the following candidate assignments of semantic value and summary table.

f: f('Madagascar') = **X**, f('this island') = Y, etc.
g: g('Madagascar') = **Y**, g('this island') = Y, etc.

Table G: Early Stage Madagascar Under KM1 & KM2

Sentence	K ex under f	K ex under g
Madagascar is Madagascar	X=X	Y=Y
This island is this island	Y=Y	Y=Y
Madagascar is this island	·	Y=Y
This island is Madagascar	·	Y=Y
Madagascar is an island	·	Y is an island
This island is an island	Y is an island	Y is an island
Madagascar was invaded	X was invaded	·
This island was invaded	·	·
Madagascar was peaceful	X was peaceful	·
This island was peaceful	·	·
Total Under KM1	**5**	**6**
Total Under KM2	**5**	**2**

The correct semantic value of 'Madagascar' at the early stage is X. KM2, unlike KM1, secures this result. In the late stage however, a fairly widespread intuition is that the semantic value of 'Madagascar' changes to Y. In the late stage, when Marco Polo picks up much more knowledge about Y, many more sentence pairs will be added of the form ⌜Madagascar is F⌝ and ⌜This island is F⌝; ⌜Madagascar is G⌝ and ⌜This island is G⌝, etc. Here is a summary table illustrating the situation at the late stage.

Table H: Late Stage Madagascar Under KM1 & KM2

Sentence	K ex under f	K ex under g
Madagascar is Madagascar	X=X	Y=Y
This island is this island	Y=Y	Y=Y
Madagascar is this island	·	Y=Y
This island is Madagascar	·	Y=Y
Madagascar is an island	·	Y is an island
This island is an island	Y is an island	Y is an island
Madagascar was invaded	X was invaded	·
This island was invaded	·	·
Madagascar was peaceful	X was peaceful	·
This island was peaceful	·	·
Madagascar is F	·	**Y is F**
This island is F	**Y is F**	**Y is F**
Madagascar is G	·	**Y is G**
This island is G	**Y is G**	**Y is G**
Madagascar is H	·	**Y is H**
This island is H	**Y is H**	**Y is H**
Total Under KM1	**8**	**12**
Total Under KM2	**8**	**5**

At the late stage the correct semantic value of 'Madagascar' is Y. This time KM1, unlike KM2, makes this prediction. Therefore, neither KM1 nor KM2 makes the intuitive predictions in **Madagascar** because neither predicts the right result at both stages. KM2 does better than KM1 at the early stage and KM1 does better than KM2 at the late stage.[16]

[16]Assuming there aren't any other terms with X as semantic value. If there are, then the semantic freedom problem will kick in and make things even worse for KM2.

Of course KM1 lines-up with intuition reasonably well when a lot of time passes and there is more knowledge about the intuitively correct and new semantic referent – the one that is playing a suitably important role in the speaker's life. Unfortunately however, as observed in §3 and §4, KM1 will not line-up with intuition if the massive amount of knowledge about Y is acquired in, say, a day. (This will essentially be a **Madagascar** version of the curiosity problem). After all, if Marco Polo learns a lot about Y in a day intuition says that he is still semantically referring to X by 'Madagascar' at the early stage. So KM1 by itself doesn't really offer any insight as to whether and why time makes a difference in the above case of semantic reference change.[17]

Might communal variants of KM make better predictions in cases of semantic change? It seems not: observations parallel to those made in §4.2 will apply here.

6 Holistic Strategies of Interpretation

Lewis (1974) formulated a version of *the problem of radical interpretation* by asking us to suppose that we have all the knowledge about some individual, but only as a physical system (e.g., how he moves, what forces he exerts on his surroundings, what light or sounds he absorbs or emits, etc.). Given all this knowledge, Lewis then asks, how might we know this individual's beliefs, desires and utterance meanings? Lewis's diagrams the problem of radical interpretation as follows: 'given ... the facts about [the individual] as a physical system, solve for the rest' (Lewis 1974, p. 331).[18]

But as Davidson (1974) and others are aware there is a prima facie challenge here. On the face of it, one cannot know the meanings of a speaker's utterances without knowing their beliefs and one cannot know the speaker's beliefs without knowing the meanings of their utterances. However, charity-based methods of interpretation are intended to meet this challenge: Davidson hopes that using charity we can solve for both belief and meaning *simultaneously*. As Glüer (2011) points out:

> For Davidson, belief and meaning are interdependent; the radical interpreter determines both belief contents and meanings simultaneously. Assignments of belief automatically fall out of

[17]One might think that it is semantically important that over time Marco Polo and his community increasingly care about Y. But it's far from clear that variation in these affective attitudes can be adequately encoded simply by observing the quantity of knowledge expressed given that knowledge is a doxastic attitude.

[18]This contrasts with the less ambitious meta-semantic project in Lewis (1983) which has served as the precedent for my discussion so far.

> assignments of meanings to sentences held true. Therefore, beliefs are as public as meanings are.

To see how this might work, take a charity-based principle that assigns meanings to sentences in a way that maximises, say, truths only. The meanings this principle assigns are also beliefs that the principle attributes to the speaker. But left as it is this principle is obviously inadequate; some of the speaker's beliefs are false and with no other constraints there is nothing stopping our charity principle from attributing beliefs that are all in fact true. Lewis (1974) therefore considers a strategy of what he calls 'successive approximation'. But even this strategy, he says, *assumes* at the beginning some 'tentative' speaker beliefs so that one may go and solve for meanings of utterances. When the initial results are in, he says, we may go on revising beliefs and meanings until a 'satisfactory balance is struck'. For various reasons Lewis (1974) thinks this strategy is inadequate and opts for a strategy under which the contents of the speaker's attitudes are filled in completely and 'no special attention is given to [the speaker's] language at [the first] step' (Lewis 1974, p. 341). At the second step however, the content of the speaker's attitudes are used to fill in the meanings of his utterances. Somewhat in the spirit of this strategy, I've been working on the following project: given facts about the speaker as a physical system and given their knowledge, solve for the meanings of their utterances. That is, I first filled in the speaker's set of knowledge and then tried to solve for utterance meaning by maximising knowledge expressed.

I should emphasise, however, that my central concerns will carry over to a strategy of using knowledge maximisation to devise a method of interpretation where knowledge and utterance meanings are somehow fixed simultaneously. Let's consider two competing interpretational packages in, say, **Leaf Raking** (though there will be an analogous issue for **Madagascar**). The first package uses assignment f to interpret the speaker's words and includes a correct specification of the speaker's knowledge. The second package uses assignment g to interpret the speaker's words and includes a correct specification of the speaker's knowledge. A simultaneous method of interpretation, however it's formulated, should not find fault with the knowledge component of these packages. This is because, by hypotheses, the knowledge component of these packages are identical and correspond to the speaker's actual knowledge. So the question arises: why should the holistic package with f in it be preferred to the holistic package with g in it? Since our two packages agree on the knowledge component, it's the differing meaning assignments that are going to have to break the tie. But as we've seen, it's very hard to see how anything in the vicinity of knowledge maximisation would explain why the f package is preferred to the g package. This suggests that knowledge maximisation cannot after all play the starring role in a holistic strategy that assigns knowledge and meaning simultaneously.

Finally, the incorrect predictions made by variants of KM weren't especially due to choosing knowledge rather than true belief as the starring propositional attitude, since the problems I raised in §3 and §4 easily extend to true belief. One may simply re-label all the above tables (e.g., change 'Total Under KM1' to 'Total Under TBM1') and use parallel cases to raise analogous problems. When the parallel cases are set up appropriately, the scores and consequently semantic predictions will remain exactly the same. Moreover, the remarks above can readily be adapted to a holistic version of the true belief strategy.

7 Conclusion

The challenge presented so far for KM is a significant one. Curiosity models made trouble for KM1. KM2 seemed like an improvement because it did better with that model. However, ultimately, KM2 confronted other problems like the semantic freedom problem and the reverse assignment problem (these also posed a challenge for KM1). One natural strategy was to switch from versions of KM that focused on the individual to versions that focused on the community, but this did not make the problems go away. And in cases of semantic change all versions of KM gave incorrect predictions at either the early or late stage. Moreover, as I showed in the previous section, the problems raised cast doubts about more holistic strategies that appeal to knowledge or true belief.

In this paper I have used candidate assignments with the intuitively correct semantic values for most terms in the set of accepted sentences (e.g., 'is', 'was', 'market', 'leaves', etc). Moreover, candidate assignments like f and g only differed with respect to what they assigned the name in question – be it 'Jones' or 'Madagascar'. When some of these assumptions are relaxed, KM is in even greater trouble. For example, we would have to rule out assignments that assigned a different piece of mathematical knowledge to every sentence. And it's a good question how various versions of KM are supposed to do this. It's striking however that even when we assume the intuitively correct semantic values for nearly all terms, KM still cannot always deliver the right result.

Of course it remains possible that something in the vicinity of KM might figure in a much more complex charity-based theory of interpretation. Clearly, I'm in no position to refute this open-ended suggestion, though the lessons of this paper will be a helpful guide to anyone wishing to develop a theory along those lines.

References

Bach, Kent. 2006. What Does It Take to Refer? *Pages 516–554 of:* Lepore, Ernest, & Smith, Barry (eds), *The Oxford Handbook of Philosophy of Language*. Oxford University Press.

Davidson, Donald. 1973. Radical Interpretation. *Dialectica*, **27**(1), 314–328.

Davidson, Donald. 1974. Belief and the Basis of Meaning. *Synthese*, **27**(July-August), 309–323.

Davidson, Donald. 1977. The Method of Truth in Metaphysics. *Midwest Studies in Philosophy*, **2**(1), 244–254.

Evans, Gareth. 1973. The Causal Theory of Names. *Aristotelian Society Supplementary Volume*, **47**(1), 187–208.

Glüer, Kathrin. 2011. *Donald Davidson: A Short Introduction*. Oxford University Press (Oup).

Kripke, Saul A. 1977. Speaker's Reference and Semantic Reference. *Pages 255–296 of:* French, Peter A., Jr, Theodore E. Uehling, & Wettstein, Howard K. (eds), *Studies in the Philosophy of Language*. University of Minnesota Press.

Lewis, David. 1983. *Philosophical Papers Volume I*. Oxford University Press.

Lewis, David K. 1974. Radical Interpretation. *Synthese*, **23**(July-August), 331–344.

McGlynn, Aidan. 2012. Interpretation and Knowledge Maximization. *Philosophical Studies*, **160**(3), 391–405.

Williamson, Timothy. 2007. *The Philosophy of Philosophy*. Wiley-Blackwell.