Fragmentation, mathematical ignorance, and the metalinguistic reply*

Adam Elga and Agustín Rayo†

May 16, 2018

Abstract

On the *coarse-grained view*, any distinctions between mental contents can be characterized using possible worlds. The coarse-grained view is sometimes thought to be refuted by straightforward counterexamples involving failures of logical omniscience. An independently motivated theory of mental fragmentation enables one avoid the counterexamples while preserving the coarse-grained view.

1 *Introduction*

How fine-grained are mental contents?

On the *coarse-grained* view, any distinctions between mental contents can be characterized using possible worlds (Stalnaker 1984, Lewis 1970, sec. V, Lewis 1975, Lewis 1994, Braddon-Mitchell and Jackson 2007). According to the simplest version of the coarse-grained view, a mental content is just a set of possible worlds.

On the *fine-grained* view, there are mental contents that differ in ways that cannot be captured using possible worlds. On one version of the fine-grained view, a mental content is a Fregean sense (Frege 1956). On another version, a mental content is a set of possible or impossible worlds (Fagin et al. 1995, Hintikka 1975). On yet other versions, mental contents are associated with sentences in an appropriate language (Field 1978, Fodor 1975).

*Thanks to Ross Cameron, David Chalmers, Jonathan Cohen, Keith DeRose, Cian Dorr, Hartry Field, Branden Fitelson, Jeremy Goodman, David Hunter, Frank Jackson, Shivaram Lingamneni, Michael Rescorla, Robert Stalnaker, Jason Stanley, Bruno Whittle, audiences at Eastern and Pacific meetings of the American Philosophical Association, the 2008 Arizona Ontology Conference, Brown University, Catholic University of Peru, CUNY, the Ohio State University, National Autonomous University of Mexico, University of Bologna, UC Berkeley, UC Riverside, UC Santa Cruz, University of Connecticut at Storrs, University of Graz, University of Leeds, University of Paris (IHPST), University of Regensburg, University of Oslo, University of Syracuse, University of Texas at Austin, and Yale University, the Corridor reading group (on two occasions), and participants in a graduate seminar session at Rutgers University and a Fall 2011 joint MIT/Princeton graduate seminar. The initial direction of this paper was enormously influenced by conversations with Andy Egan.

†adame@princeton.edu, arayo@mit.edu

The coarse-grained view is sometimes thought to be refuted by straightforward difficulties involving logical or mathematical ignorance. In this paper we will argue that this is not so. We will show that a friend of coarse-grained contents has the resources to address the difficulties. Of course, fending off difficulties is not the same as showing that the coarse-grained view is correct. An overall assessment of the strength of the view is beyond the scope of this paper.

2 *The coarse-grained view*

In this section we will describe a simple example of a coarse-grained view of mental content. It is based on a coarse-grained notion of information according to which a batch of information is just a set of possible worlds—the worlds at which that information obtains.

How does the possession of a given batch of information relate to states of belief, and to belief attributions? And what, exactly, is the connection between information possession and rational action? Here is a SIMPLE PLAN for answering these questions:

SIMPLE STATE A subject's belief state is the information she possesses. (Since a batch of information is taken to be a set of possible worlds, this means that the subject's belief state is a set of worlds.)

SIMPLE ATTRIBUTION A subject believes that p if and only if every world in her belief state is such that p.

SIMPLE ACTION An action is predicted for a subject if and only if it tends to bring about her desires at worlds compatible with the information she possesses.

Suppose, for example, that a soldier walking a minefield possesses the information that there are no mines by the river (Braddon-Mitchell and Jackson 2007). Then according to SIMPLE STATE the soldier's belief state is a set of worlds at which there are no mines by the river, and according to SIMPLE ATTRIBUTION, the soldier believes that there are no mines by the river. Finally, suppose that the soldier also possesses the information that the high ground is mined and that her options are traveling by the river or traveling on high ground. Then SIMPLE ACTION predicts that she will travel by the river (assuming she desires not to step on a mine).

A more sophisticated version of SIMPLE PLAN would replace each of the plan's three theses with probabilistic analogues.[1] Here we will mostly focus

[1] For instance:

on non-probabilistic versions of SIMPLE PLAN, but it is worth keeping in mind that our arguments can be adapted to fit the probabilistic case.

Some philosophers have argued that a theory of mind based on SIMPLE PLAN would be hopeless (Field 1986b, Speaks 2006, Williamson 2016). We agree. SIMPLE PLAN faces insurmountable difficulties, to which we now turn.

3 *Problem 1: metaphysical omniscience*

It is an immediate consequence of SIMPLE STATE and SIMPLE ATTRIBUTION that belief ascriptions are closed under metaphysical entailment: if every world at which ϕ is true is also a world at which ψ is true, then $\ulcorner S$ believes that $\phi\urcorner$ can only be true if $\ulcorner S$ believes that $\psi\urcorner$ is.[2]

An unwelcome consequence of this result is that SIMPLE PLAN entails that all subjects believe all logical and mathematical truths, and that their beliefs respect all logical entailments. Such results are wildly implausible (Field 1986a, 428, Soames 1985, 59–60, Soames 1987, 48–50, Speaks 2006, 446). This *problem of logical omniscience* is our present concern.

3.1 *A simple metalinguistic reply*

The coarse-grained theorist might follow Stalnaker (1984) and Braddon-Mitchell and Jackson (2007, 200–201), and reply that mathematical and logical uncertainty can be partly metalinguistic.[3] For example, she might posit that

---

SIMPLE STATE* A subject's belief state is her credence function—her assignment of subjective probabilities to propositions, understood as sets of possible worlds.

SIMPLE ATTRIBUTION* A subject is confident that p if and only if she assigns high enough credence to the set of worlds in which p.

SIMPLE ACTION* An action is predicted for the subject if and only if no rival action is assigned greater expected value by subject's credence function (given the subject's values).

[2]Proof: suppose that ϕ metaphysically entails ψ and that $\ulcorner S$ believes that $\phi\urcorner$ is true. Then by SIMPLE ATTRIBUTION, ϕ is true at every world in S's belief state, and hence ψ is true at every world in S's belief state. So by SIMPLE ATTRIBUTION, $\ulcorner S$ believes that $\psi\urcorner$ is true.

[3]Stalnaker thinks that "Under certain conditions, the content of an assertion is not the proposition determined by the ordinary semantic rules, but instead the diagonal proposition of the propositional concept determined." (Stalnaker 1987a, 124) and argues that something similar happens in the case of belief attributions. Braddon-Mitchell and Jackson (2007, 201) endorses a similar idea: "When Jones discovers that a triangle she already knew to be equiangular is also and necessarily equilateral, she does indeed discover something,

uncertainty about whether $12 \times 18 = 216$ can derive in part from uncertainty about the following metalinguistic claim:

C The sentence "$12 \times 18 = 216$" is true.

Since C is contingent, the coarse-grained theorist might hope to avoid the implication that everyone believes C. And she might hope to analyze mathematical uncertainty in terms of uncertainty about claims such as C.

Here is a problem for this metalinguistic reply.[4] In representing a subject's mental state, it is not enough to do justice to what information she lacks. We must also do justice to the information that she possesses. Even if an ordinary subject is ignorant of the fact that $18 \times 5 = 90$, she can be expected to know the meanings of mathematical vocabulary and the compositional rules of mathematical language. So she will know—and hence believe—basic metalinguistic claims such as that "18" refers to 18. But these basic metalinguistic claims together entail that

> "$12 \times 18 = 216$" is true if and only if $12 \times 18 = 216$,

and hence entail that C is true.

Now, according to SIMPLE STATE and SIMPLE ATTRIBUTION, one believes whatever is entailed by propositions one believes. So SIMPLE STATE and SIMPLE ATTRIBUTION entail that everyone believes true metalinguistic claims such as C. So analyzing mathematical uncertainty in terms of metalinguistic uncertainty cannot be enough to defuse the problem of logical omniscience.

---

but arguably not about how things are but about how it is that two different sentences constructed out of interestingly different materials represents the very same way things are."

[4]Problems of this kind have been raised by Field (1978, 34–35), Field (1986b), Soames (2009), Speaks (2006, 448–450), Williamson (2016, n. 1) and Bacon (2017). Here is how Williamson presses the point:

> One problem for his view is that for any formula A in standard mathematical notation, the biconditional True($\ulcorner A \urcorner$) $\leftrightarrow A$ will follow logically from known axioms of a standard compositional theory of truth for the mathematical language. Since on Stalnaker's view our knowledge is closed under such logical consequence, it implies that we already know the biconditional, so the metalinguistic claim is equivalent for us to the original mathematical claim. Thus semantic ascent to the metalinguistic level only postpones the problem. Williamson (2016, footnote 1)

4 Problem 2: actions arising from mathematical ignorance

An additional problem for SIMPLE PLAN arises not from belief attributions, but rather from the relationship between information possession and dispositions to act.

The problem is best explained with an example.[5] Suppose that Alice and Bob have the same desires but that they each have a false belief about 12×18: Alice thinks that it equals 214, Bob thinks that it equals 218. They are remodeling their bathroom, and the plan calls for a grid of 12 tiles by 18 tiles. When shopping for the required tiles, Alice will be disposed to buy a different number of tiles than Bob.

Such subjects are clearly possible, but it is difficult to see how a friend of SIMPLE PLAN can capture Alice and Bob's mathematical ignorance in a way that predicts the correct tile-shopping behavior.

To see why, notice that *no* possible worlds are compatible with 12×18 being equal to 214 (as Alice thinks) or 218 (as Bob thinks). So in the absence of a special story, the coarse grained theorist must represent each of Alice and Bob's belief states by the very same set: the empty set. But according to SIMPLE ACTION, differences in predicted behavior between subjects only derive from differences in the possible worlds compatible with each agent's beliefs (when desires are held fixed). So it seems that the coarse grained theorist cannot account for the difference between Alice and Bob's tile-shopping behavior.

A friend of SIMPLE PLAN might insist that a suitable difference between Alice and Bob's belief states are ready to hand: Alice believes, for example, that 214 tiles are required to for the bathroom, while Bob believes that 218 are required. There is no problem representing *that* difference in terms of sets of possible worlds, since the set of worlds in which 214 tiles are required is distinct from the set of worlds in which 218 are required. Furthermore, that difference seems to produce just the right predicted actions according to

[5] We take this example to capture the key point of an important example introduced in (Field 1986b):

> if I offer someone who doesn't know much mathematics $1000 for an example of a plane map that requires more than four colors to color (according to the usual coloring conventions), he will behave very differently than he would if I had offered $1000 for a trisection of a Euclidean 60 degree angle by straight edge and compass; to explain this, I need to attribute different beliefs and desires to him in the different cases, and it is prima facie difficult to see how I can do this in a relevant way if the desire to do one impossible task is identified with the desire to do any other impossible task.

SIMPLE ACTION.

However, this proposal is vulnerable to a difficulty analogous to the one raised above for the metalinguistic reply. As before, it is not enough for the coarse grained theorist to do justice to the information that Alice and Bob lack. He must also do justice to the information that they possess. In particular, Alice and Bob each believe that the bathroom requires 12×18 tiles. And any world compatible this fact is a word at which the bathroom requires 216 tiles. These considerations put pressure on the coarse-grained theorist to represent Alice and Bob's belief states by sets containing only worlds in which the bathroom requires 216 tiles. But doing so will not account for either Alice or Bob's tile-shopping behavior.

So SIMPLE PLAN does not supply the coarse-grained theorist with an adequate way to represent a failure of logical omniscience.

5 *The state of play*

Problems 1 and 2 should not be understood as the complaint that coarse-grained theorists have failed to produce a general theory of ordinary-language belief attributions or a general theory of predicted action. That demand would be unreasonable, since it is so difficult to give such theories in any framework whatsoever. Rather, the challenge is that certain deep structural features of the coarse-grained framework seem to immediately rule out any such theory.

We will answer this challenge by giving up on SIMPLE PLAN but retaining a coarse-grained view of mental contents.

Our motivating idea is that the across-the-board notion of information possession should be replaced with a notion of information accessible *relative to a condition*. We will motivate this picture and then use it to address Problems 1 and 2.

6 *Information access*

Why replace the across-the-board notion of information possession with a relative notion of accessible information? Because information can be represented in a way that makes it accessible for some purposes, but inaccessible for others (Stalnaker 1991, 437–438).

Consider a pair of crossword-puzzle solvers trying to fill in the blanks below to complete a word of English:

_ _ _ _ M T

The first puzzlist fills in just the right letters. The second scratches his head and leaves the puzzle blank. Suppose further that each puzzlist knows that "dreamt" is a word of English, and knows how to spell it. Indeed, each puzzlist realizes from the start that "dreamt" is a word of English that ends in MT.

So why is one puzzlist disposed to fill in the blanks with DREA, while the other is disposed to gnash his teeth, curse, and fill in nothing?

We suggest that both puzzlists possess the information they need to fill in the blanks, but that the conditions relative to which they have access to this information are different. Let D be the set of worlds in which *dreamt* is a word of English spelled D-R-E-A-M-T. Both puzzlists have access to D for the purpose of using the word "dreamt" in a written essay. And they both have access to D for the purpose of answering the question "Is 'dreamt' a word of English ending in MT?".

But for the purpose of filling in the blanks in "_ _ _ _ M T", only the first puzzlist has access to D.

So if we'd like to represent the difference between the two puzzlists, our representation of each of them should not attempt to specify what information he possesses, period. It should instead say what information he or she has access to for what purposes.[6]

For example, we might represent the informational state of the struggling puzzlist with the following sort of table, which we shall call an *access table*:

Choice condition	*Accessible information*
dreamt salient	$I_{!}$
dreamt not salient	$I_{?}$
[more conditions]	[information accessible relative to those conditions]

[6] In an illuminating paper on the role of the organization of memory in human reasoning, Cherniak (1983, 166) notes that creatures with mental organizations anything like ours constantly face a hard computational problem: quickly accessing memories relevant to their current situation. Cherniak convincingly argues that there is therefore a practical need for a small short-term memory store that supports fast—but not exhaustive—searching and consistency checking. It is to be expected that the heuristics underlying such searches will operate differently in different circumstances, and hence that different information will be accessible in different circumstances. Indeed, the necessity for heuristics that narrow memory searches was already recognized by Hume: "as the production of all the ideas to which [a] name may be applied, is in most cases impossible, we abridge that work by a more partial consideration, and find but few inconveniences to arise in our reasoning from that abridgement" (Hume 1738, 21, as cited in Cherniak 1983, 176).

Here $I_!$ and $I_?$ are sets of possible worlds representing batches of (coarse-grained) information. $I_!$ contains only worlds in which *dreamt* is a word of English spelled D-R-E-A-M-T. $I_?$ contains other worlds as well.

The table reflects that the puzzlist's dispositions factor into two natural components: one component associated with situations in which the word "dreamt" has been made salient, and another in which it has not.[7]

7 *The fragmentationist plan*

Once a relativised notion of information possession is in place, it is straightforward to amend SIMPLE STATE and SIMPLE ACTION appropriately:

FRAGMENTED STATE A subject's belief state is ~~the information she possesses~~ **her access table**.

FRAGMENTED ACTION An action is predicted for a subject if and only if it tends to bring about her desires at worlds compatible with the information she possesses **relative to her current choice condition**.

We also need a replacement for SIMPLE ATTRIBUTION. Here is the original version:

SIMPLE ATTRIBUTION A subject believes that p if and only if every world in her belief state is such that p.

In proposing a replacement, our goal is modest. We will not offer a full defense of any particular conception of belief. We will not provide a compositional semantics for belief attributions.[8] We aim only to show that

[7] The suggestion in Stalnaker (1984) that logical omniscience failures can be understood in terms of fragmented belief states was the core motivation for the present model. Braddon-Mitchell and Jackson (2007, 199–200) also uses fragmented coarse-grained belief states to accommodate failures of logical omniscience. Yalcin (2008, Ch. 3), Yalcin (2015), and Yalcin (2016) develop that same suggestion, proposing an elegant model on which all-or-nothing belief is relative to questions, understood as partitions of logical space. The treatment of logical omniscience failures in those works uses privileged partitions to represent which propositions are accessible to an agent, and so differs from the present treatment. (See especially Yalcin (2016, n. 26).) Egan (2008) endorses a treatment of fragmented credences similar to the present one and interestingly suggests that mental fragmentation might be practically indispensable for agents with perceptual belief forming mechanisms anything like human ones—mechanisms that are less than perfectly reliable but which nevertheless produce immediate belief in certain circumstances.

[8] One way to pursue this project in the present context would be to build on the picture of belief ascriptions articulated in Stalnaker (1988, 1987b).

Problems 1 and 2 are not, on their own, enough to rule out coarse-grained conceptions of mental content. To achieve that aim it is enough to articulate a respectable conception of belief that is based on a coarse-grained conception of mental contents and has principled answers to Problems 1 and 2.

The respectable conception of belief we will consider is a variation of the dispositional analysis of belief set forth in Schwitzgebel (2002, 253): "To believe that P [...] is nothing more than to match to an appropriate degree and in appropriate respects the dispositional stereotype for believing that P."

What is the dispositional stereotype associated with a given belief? It is a cluster of dispositions to act and react "in patterns that ordinary people would regard as characteristic of having that [belief]" (Schwitzgebel 2013, 75). This is only a rough description, but as far as we know no one has a general reductive analysis to offer here. (Indeed, we doubt that any simple, systematic story will suffice.) Schwitzgebel does give some helpful examples:

> Consider a favorite belief of philosophers: the belief that there is beer in the fridge. Some of the dispositions associated with this belief include: the disposition to say, in appropriate circumstances, sentences like 'There's beer in my fridge'; the disposition to look in the fridge if one wants a beer; a readiness to offer beer to a thirsty guest; the disposition to utter silently to oneself, in appropriate contexts, 'There's beer in my fridge'; [...] and so forth."[9] (Schwitzgebel 2002, 251)

We would like to propose a slight modification of this account of belief:[10]

FRAGMENTED ATTRIBUTION To have a given belief is to have an access table that predicts a sufficient portion of the dispositional stereotype associated with that belief.

[9]In the list of stereotypical dispositions quoted above, Schwitzgebel also includes "an aptness to feel surprise should one go to the fridge and find no beer". We leave it out here because we wish to remain neutral on the question of whether dispositional stereotypes include not just dispositions to act, but also dispositions to feel emotions and have experiences. Schwitzgebel's list also includes "the disposition to draw conclusions entailed by the proposition that there is beer in the fridge". We exclude this disposition because the relevant notion of entailment would need to be carefully constrained, and describing such constraints is outside of the present focus.

[10]This proposal is far from a full analysis of belief attributions. Our goal here is less ambitious: a first-approximation account of belief ascriptions that makes room for failures of logical omniscience. Though our account will not be adequate in general (among other limitations, the account does not even hope to handle difficulties raised in Soames (2009)), we hope it will show that there is no essential tension between coarse-grained accounts of content and failures of logical omniscience.

This delivers a coarse-grained account of belief because it entails that having a belief amounts to having access to particular batches of coarse-grained information relative to particular circumstances.

For a subject with fixed desires, FRAGMENTED ATTRIBUTION tells us that having a particular belief means having access to appropriate information. But it does not associate a *single* piece of information with each belief and say that having the belief requires having access to that piece of information in all conditions. Rather, FRAGMENTATED ATTRIBUTION says that to have a given belief a subject must have access to *various* batches of information in various conditions. Which batches and which conditions? Exactly the ones that lead to the dispositions stereotypical of that belief.

Here is an example. Recall the crossword puzzlist who is unable to fill in "_ _ _ _ M T" even though he believes that *dreamt* is a word of English spelled D-R-E-A-M-T. What dispositional stereotype corresponds to this belief? A typical answer might include patterns of dispositions such as the following. When assigned the task of writing an essay about his dreams on the previous night and using the word *dreamt*, he is disposed to spell it correctly. When shown DREAMT and asked whether it is a properly spelled word of English, he is disposed to answer "yes". And so on. So according to FRAGMENTED ATTRIBUTION, for this subject to believe that *dreamt* is a word of English spelled D-R-E-A-M-T is for him to have an access table that predicts a sufficient portion of the above dispositions.

Notice that FRAGMENTED ATTRIBUTION allows for the practice of belief attribution to "play favorites" by counting only certain elicitation conditions as relevant to a belief ascription.

For example, recall that D is the set of worlds in which *dreamt* is a word of English spelled D-R-E-A-M-T. In order for our puzzlist to count as believing that *dreamt* is a word of English spelled D-R-E-A-M-T, he need not have access to D relative to all conditions. In particular, he need not have access to D for the purpose of filling the blanks in "_ _ _ _ M T".

Here ends our explanation of FRAGMENTED PLAN, which we claimed above would help address Problems 1 and 2 from §§3–4.[11] It is time to

[11] It is worth noting that one could also formulate a *probabilistic* version of FRAGMENTED PLAN, according to which an access table assigns to each elicitation condition a *probability function* rather than a set of worlds. An important advantage of using probability functions is that they allow us to take advantage of Bayesian decision theory, and replace FRAGMENTED ACTION with a fragmented variant of decision theory which allows for agents who have imperfect access to their information:

FRAGMENTED DECISION THEORY An option is predicted for a subject in a given elicitation condition if and only if no rival option is assigned greater expected value by the probaddress those problems.

8 *Response to Problem 1*

Problem 1 is based on the observation that SIMPLE STATE and SIMPLE ATTRIBUTION require that belief ascriptions are closed under metaphysical entailment. FRAGMENTED PLAN avoids this difficulty because FRAGMENTED STATE and FRAGMENTED ATTRIBUTION require no such thing.

To see why, focus again on the case of the crossword puzzlist from the previous section. Consider the following belief attributions:

1. The puzzlist believes that *dreamt* is an English word spelled D-R-E-A-M-T.
2. The puzzlist believes that to solve the puzzle, what is needed is an English word ending in MT.
3. The puzzlist believes that *dreamt* solves the puzzle.

The contents attributed[12] in (1) and (2) together metaphysically entail the content attributed in (3). But according to FRAGMENTED ATTRIBUTION, (1) and (2) are true and (3) is false when the puzzlist's access table is as in §6.[13]

9 *Response to Problem 2*

Recall Alice and Bob remodeling a bathroom floor. They each realize that the floor requires a grid of 12 by 18 tiles. But the two of them have different false beliefs about the value of 12×18. As a result, they are disposed to buy different numbers of tiles.

Problem 2 is based on the claim that a coarse-grained picture based on SIMPLE STATE and SIMPLE ACTION cannot satisfactorily model the mental

ability function in the row associated with that elicitation condition in the subject's access table.

[12]By the "content attributed" in a belief attribution we mean the content of the sentential complement of "believes that" in that attribution.

[13]The key observation is that (1) and (3) are plausibly associated with different dispositional stereotypes. As we suggested above, the stereotype corresponding to (1) need not include the disposition to correctly fill in the blanks in "__ __ __ __ M T". In contrast, the stereotype corresponding to (3) would plausibly include such a disposition. But when the puzzlist's access table is as in §6, he has access to the information that *dreamt* is a word of English spelled D-R-E-A-M-T only relative to conditions in which *dreamt* is salient. So FRAGMENTED ATTRIBUTION counts him as believing (1) but not as believing (3).

states of Alice and Bob. The basic difficulty is that there is a tension in Alice and Bob's beliefs. Alice believes both that the bathroom requires 12×18 tiles and that it requires 214 tiles. Bob believes both that the bathroom requires 12×18 tiles and that it requires 218 tiles. As a result, there are *no* possible worlds compatible with every aspect of Alice's beliefs, or every aspect of Bob's beliefs. So SIMPLE PLAN requires that their belief states are both represented by the empty set, which means that SIMPLE ACTION won't predict the important behavioral differences between Alice and Bob.

Adopting FRAGMENTED PLAN avoids this difficulty because it removes the pressure to represent Alice and Bob's belief states by the empty set. For example, we might represent Alice's belief state using an access table of the following sort:

Choice condition	*Accessible information*
Asked questions that test basic understanding of arithmetic	I_{und}
Asked "Is it the case that $12 \times 18 = 214$?"	I_{lim}
[more conditions]	[information accessible relative to those conditions]

Here I_{und} is a set of worlds in which all mathematical vocabulary items have their standard meanings (the numeral "12" refers to the number 12, "$\times$" has multiplication as its semantic value, and so on). At every such world, the sentence "$12 \times 18 = 214$" is false.

In contrast, I_{lim} is a set of worlds in which the sentence "$12 \times 18 = 214$" is true. The worlds in I_{lim} involve various oddities: in some of them, the numeral "12" has a nonstandard meaning; in others the symbol "$\times$" does; in others the language has nonstandard compositional rules.

This table does justice to two facts about Alice. Its first row captures the fact that Alice *understands arithmetical vocabulary*. For example, if she were to be tested on whether "$\times$" expresses multiplication, or whether $\ulcorner 1 \times n = n \urcorner$ is true (for n a small enough numeral), she would pass the test.

Its second row captures the fact that Alice *has a cognitive limitation*: she wrongly believes that $12 \times 18 = 214$. For example, she would answer affirmatively if asked "Is it the case that $12 \times 18 = 214$?".

Bob's belief state might be represented with a table whose first row is the same as Alice's but whose second row contains a set of worlds in which the sentence "$12 \times 18 = 218$" is true. The resulting tables predict the desired tile-buying behavior for Alice and Bob, but appeal only to course grained content (since the batch of information at each row consists of a set of possible worlds). This answers Problem 2.

10 Objections and Replies

In §7, we articulated a conception of belief based on FRAGMENTED ATTRIBUTION and, in §§8–9, we used it to address Problems 1 and 2.

Our response to Problem 2 suggests a strategy for representing the mathematical beliefs of non-omniscient agents on which such agents are fragmented: they have access to different information relative to different elicitation conditions. We will call it the *fragmentation account*. In this section we consider objections to the fragmentation account and offer some replies.

10.1 *Objection: bringing in language is* ad hoc

Objection: On the fragmentation account, linguistic information figures heavily in the representation of mathematical belief. For example, part of what makes it true that Alice believes that $12 \times 18 = 218$ is that her access table has in an appropriate row a set of worlds at which the *sentence* "$12 \times 18 = 218$" is true.

Bringing in language in this way is *ad hoc*. A major shortcoming of the coarse-grained view of mental content is that it cannot distinguish between, for example, the content that $12 \times 18 = 214$ and the content that $12 \times 18 = 218$. The fragmentation account attempts to compensate for this shortcoming by associating contingent linguistic information with mathematical beliefs and using the linguistic information to play the role that mathematical content should be playing. But such a move is artificial and strained. It cannot be accepted unless it comes with some sort of independent motivation and none has been given.

Reply: The fragmentationist account is based on FRAGMENTED ATTRIBUTION, which tells us that having a belief is a matter of having an access table that predicts a sufficient portion of the dispositional stereotype associated with that belief.

The dispositional stereotype associated with a belief will often include linguistic dispositions—ones whose manifestation involves the deployment of linguistic information. Think back to the belief that there is beer in the fridge. Plausibly, its dispositional stereotype includes "the disposition to say, in appropriate circumstances, sentences like 'There's beer in my fridge' " (Schwitzgebel 2002, 251). Such a disposition requires the subject to possess linguistic information, since she must know the meaning of sentences like "there is beer in my fridge".

Of course, the dispositional stereotype associated with the belief that there is beer in the fridge also includes many non-linguistic dispositions. And the linguistic dispositions need be no more central or prevalent than the non-linguistic ones. But when it comes to complicated mathematical beliefs, linguistic dispositions play a more central role.

Consider, for example, the belief that every even number greater than 2 is the sum of two primes. What dispositions will its stereotype include? It is easy to think of stereotypical dispositions that make essential use of linguistic information. Such dispositions include the disposition to utter (in appropriate circumstances) sentences like "every even number greater than 2 is the sum of two primes", and the disposition to make inferences that rely on the truth of such sentences.[14] But it is not so easy to think of many stereotypical dispositions that do not rely on linguistic information. And the more complex and abstract the mathematical belief, the harder it will be.

So it is a natural consequence of a dispositional account of belief that complex mathematical beliefs will be especially tightly linked to linguistic dispositions. The fragmented view inherits this tight linkage naturally, and does not posit *ad hoc* "special cases" for mathematical belief.

10.2 Objection: subject matter

Objection: On the fragmentation account, Alice counts as believing that $12 \times 18 = 218$ because she has access to the information that the *sentence* "$12 \times 18 = 218$" is true, relative to suitable elicitation conditions. That's just wrong. Surely, Alice's belief is about mathematics (the value of the product of two numbers), not semantics (the meanings of certain words). More generally, the fragmentationist account fails to deliver a plausible story about the subject matter of mathematical beliefs.

Reply: We agree that mathematical beliefs do not in general have language as their subject matter.[15] We deny that the fragmentation account delivers such a result.

The fragmentation acount does not entail that the subject matter of typical mathematical beliefs includes linguistic information. Rather, it entails that

[14] Compare to Schwitzgebel (2013, 89).

[15] Obvious exceptions include, for example, the belief that the number of languages Jones speaks is prime.

that the *dispositional stereotypes* that are associated with typical mathematical beliefs include linguistic dispositions, and therefore require access to linguistic information. But the dispositions stereotypical of a belief can require some information even if that information is not the subject matter of that belief.

For example, someone who believes that there is beer in the fridge will typically be disposed to utter, in appropriate circumstances, sentences like "there is bridge in the fridge". So according to FRAGMENTED ATTRIBUTION, someone who believes that there is beer in the fridge will typically have access to relevant linguistic information. But this has no tendency to show that the subject matter of the belief is language. In this case, the subject matter is whether there is beer in the fridge.

So what is the subject matter of a mathematical belief? As in the case of non-mathematical beliefs, the subject matter of a mathematical belief is the feature of the world on which the truth or falsity of the belief depends. So, for example, the subject matter of the belief that $12 \times 18 = 218$ is whether $12 \times 18 = 218$.

An opponent might wish to press the point further and ask: "What settles whether $12 \times 18 = 218$?" That is a difficult question in the philosophy of mathematics which we can't hope to settle here. We would like to point out, however, that the proposal we are defending does not depend on the answer one gives. For example, on one view what settles whether $12 \times 18 = 218$ is whether certain non-spatiotemporal abstract objects stand in appropriate arithmetical relations. That view is compatible with thinking that the dispositions stereotypical of the belief that $12 \times 18 = 218$ include answering certain linguistically posed questions in a distinctive way.

10.3 Objection: Linguistic Ignorance

Objection: Recall that Alice believes that $12 \times 18 = 218$. The fragmentation view represents her mental state using an access table that involves linguistically deviant possible worlds: worlds in which mathematical vocabulary items have nonstandard meanings (for example, "$\times$" might denote a function different than the multiplication function).

Linguistically deviant worlds do have a place in representing certain types of mathematical confusion or uncertainty. Consider, for example, a child who is just starting to learn arithmetic and doesn't fully understand arithmetical vocabulary. It makes sense to represent her belief states using linguistically deviant worlds. In contrast, Alice's uncertainty is nothing like *that*. Alice is wrong about the value of 12×18, but there's nothing wrong with her understanding of the symbol "$\times$". Like most of us, Alice knows full well what "$\times$" means, so deviant worlds in which "$\times$" stands for some function other than multiplication have no place in the representation of her mathematical beliefs. Her beliefs rule out such worlds from the start.

This means that the fragmentation account is wrong to appeal to linguistically deviant worlds in representing Alice's belief state. What's worse, the case of Alice is just the tip of the iceberg. The fragmentation account appeals to linguistically deviant worlds to represent the belief states of every non-ideal thinker, even the most sophisticated mathematicians in the world. And it is absurd to think that the most sophisticated mathematicians in the world are at all confused or ignorant about the meanings of symbols such as "$\times$".

Reply: Consider Silvia, a sophisticated mathematician who nevertheless is ignorant of some multiplication fact. The objection is that the fragmentation account entails that Silvia is confused or ignorant about the meanings of symbols such as "$\times$". We reply that the fragmentation account entails no such thing. Indeed it entails the opposite—that Silvia has just the sorts of beliefs that are associated with excellent understanding of the meaning of "$\times$". It entails, for example, that Silvia believes that for all n, $n \times 0 = 0$, and that she believes that for all n and k, $n \times (k+1) = (n \times k) + n$. This is because Silvia's access table predicts a sufficient portion of the dispositions stereotypical of each of these beliefs.[16]

10.4 Objection: shared beliefs

Objection: On the fragmentation account, it is hard to accommodate the fact that people who don't share a language can share a mathematical belief. For example, well-educated monolingual English geometers believe that the square of a triangle's hypotenuse is equal to the sum of the squares of the triangle's other sides, and ancient Greek geometers believed the vary same thing. So there is a mathematical belief modern and ancient geometers share. But their linguistic dispositions must differ, since they don't share a language. So on the fragmentation account, the two groups of geometers

[16] The above beliefs are the sorts of beliefs associated with excellent understanding of "$\times$", on any everyday sense of "excellent understanding". One might introduce a more demanding notion—call it "perfect understanding of a symbol"—requiring that one's grasp of a symbol not even indirectly contain any incoherence or latent confusion. We grant that the fragmentation account entails that no human's understanding of "$\times$" is perfect in that sense. But that is an entirely plausible conclusion, since no human is infallible about arithmetic.

don’t share a belief after all.[17]

Reply: The first thing to note that the objection is not specific to mathematical belief. It can be pressed against any dispositional account of belief that appeals to linguistic dispositions. For example, one might observe that modern Americans and ancient Egyptians share the belief that the pyramids are stable. One might object that a dispositionalist account of belief cannot do justice to this fact because modern Americans and ancient Egyptians don’t share a language.

The dispositionalist has the resources to answer this objection. She might claim, for example, that the relevant stereotype includes not the disposition to utter, in appropriate circumstances, suitable *English* sentences, but the disposition to utter, in appropriate circumstances, suitable sentences *of a language one understands*. Or she might claim that different stereotypes are appropriate for different subjects in different circumstances.[18] As far as the present discussion is concerned, nothing turns on how one chooses to address the issue. And however one addresses it in the non-mathematical case, a similar move will be available in the mathematical case.

10.5 The fragmentation account is the metalinguistic reply in disguise

Objection: The fragmentation account is just a variant of the idea that mathematical uncertainty boils down to uncertainty about linguistic matters. In other words: we’re back to the metalinguistic reply (§3.1). And the metalinguistic reply is implausible on its face, since mathematical ignorance is ignorance of mathematics, not language.

Moreover, as noted in §3.1, there is a powerful argument for thinking that the metalinguistic reply fails on its own terms. That argument applies to the fragmentation account just as it applies to simpler versions of the metalinguistic reply. For everyone must grant that a typical subject knows the meanings of mathematical vocabulary. And in the context of a coarse-grained view, such as the fragmentation account, one must conclude from this that subjects are certain about the truth of every sentence expressing an arithmetical truth, and therefore certain about all of mathematics.

Reply: The fragmentation account certainly agrees with the metalinguistic reply that, in a large range of cases, mathematical uncertainty is correlated with linguistic uncertainty. However, it is no part of the fragmentation

[17] For objections in this spirit, see Moore (1994, 97–100) and Nuffer (2009, §4).

[18] Schwitzgebel himself would appear to favor a third option. See Schwitzgebel (2002, 253–254, 265).

account that the *subject matter* of mathematical belief is linguistic (§10.2). So the account is not implausible on its face for reasons of mischaracterizing the subject matter of mathematical uncertainty.

In addtion, FRAGMENTED ATTRIBUTION does not entail that belief attributions are closed under metaphysical entailment (§8). This makes the fragmentation account consistent with thinking that a typical subject can know the meanings of the mathematical vocabulary while being uncertain about many mathematical matters (§9).[19]

11 Conclusion

The problem of logical omniscience raised by Problems 1 and 2 is often thought fatal to the coarse-grained view of mental content. But it is not: an appeal to mental fragmentation saves the day. Of course, other attacks based on the problem of logical omniscience are surely waiting in the wings.

References

Andrew Bacon. Response. Presented at Pacific APA session on Elga and Rayo's "Fragmented Decision Theory", April 2017.

David Braddon-Mitchell and Frank Jackson. *The Philosophy of Mind and Cognition, Second Edition*. Blackwell, 2007.

Christopher Cherniak. Rationality and the structure of human memory. *Synthese*, 57(2):163–186, 1983.

Andy Egan. Seeing and believing: Perception, belief formation and the divided mind. *Philosophical Studies*, 140(1):47–63, 2008.

Ronald Fagin, Joseph Y. Halpern, and Moshe Y. Vardi. A nonstandard approach to the logical omniscience problem. *Artificial Intelligence*, 79(2):203 – 240, 1995. ISSN 0004-3702. doi: https://doi.org/10.1016/0004-3702(94)00060-3. URL http://www.sciencedirect.com/science/article/pii/0004370294000603.

Hartry Field. Mental representation. *Erkenntnis*, 13:9–61, July 1978.

[19] It is worth keeping in mind that Stalnaker never took the metalinguistic reply, considered in isolation, to suffice for a solution to the problem of logical omniscience. As developed in Stalnaker (1984), the metalinguistic reply is part of a more complex view that also includes a reliance on fragmented belief states. A helpful critical discussion appears in Speaks (2006, 449–450).

Hartry Field. Critical notice: Stalnaker, Robert "Inquiry". *Philosophy of Science*, 53(3):425–448, 1986a.

Hartry Field. Stalnaker on intentionality: On Robert Stalnaker's "Inquiry". *Pacific Philosophical Quarterly*, 67:98–112, April 1986b.

Jerry A. Fodor. *The Language of Thought*. Harvard University Press, 1975.

Gottlob Frege. The thought: A logical inquiry. *Mind*, 65(259):289–311, 1956.

Jaakko Hintikka. Impossible possible worlds vindicated. *Journal of Philosophical Logic*, 4(4):475–484, 1975.

David Hume. *A treatise of human nature*. Clarendon Press, Oxford, 1738. Reprint 1966.

David Lewis. General semantics. *Synthese*, 22(1-2):18–67, 1970.

David Lewis. Languages and language. In Keith Gunderson, editor, *Minnesota Studies in the Philosophy of Science*, pages 3–35. University of Minnesota Press, 1975.

David Lewis. Reduction of mind. In Samuel Guttenplan, editor, *Companion to the Philosophy of Mind*, pages 412–431. Blackwell, 1994.

Robert C. Moore. *Logic and Representation*. Center for the Study of Language and Information, 1994.

Gerhard Nuffer. Stalnaker on mathematical information. *Southern Journal of Philosophy*, 47(2):187–204, 2009.

Eric Schwitzgebel. A phenomenal, dispositional account of belief. *Noûs*, 36 (2):249–275, 2002.

Eric Schwitzgebel. *A Dispositional Approach to Attitudes: Thinking Outside of the Belief Box*, pages 75–99. Palgrave Macmillan UK, London, 2013. ISBN 978-1-137-02652-1. doi: 10.1057/9781137026521_5. URL `http://dx.doi.org/10.1057/9781137026521_5`.

Scott Soames. Lost innocence. *Linguistics and Philosophy*, 8(1):59–71, 1985.

Scott Soames. Direct reference, propositional attitudes, and semantic content. *Philosophical Topics*, 15(1):47–87, 1987.

Scott Soames. Understanding assertion. In Scott Soames, editor, *Philosophical Essays, Volume 2: The Philosophical Significance of Language*, pages 211–242. Princeton University Press, 2009.

Jeff Speaks. Is mental content prior to linguistic meaning?: Stalnaker on intentionality. *Noûs*, 40(3):428–467, 2006.

Robert Stalnaker. *Inquiry*. MIT Press, Cambridge, Mass., 1984.

Robert Stalnaker. Semantics for belief. *Philosophical Topics*, 15(1):177–190, 1987a.

Robert Stalnaker. Belief attribution and context. In Robert H. Grimm and D. D. Merrill, editors, *Contents of Thought*, pages 140–156. University of Arizona Press, 1988.

Robert Stalnaker. The problem of logical omniscience. *Synthese*, 89(3):425–440, 1991.

Robert C. Stalnaker. Semantics for belief. *Philosophical Topics*, 15, 1987b. In Stalnaker (1999), pp. 117–129.

Robert C. Stalnaker. *Context and Content*. Oxford University Press, Oxford, 1999.

Timothy Williamson. Absolute provability and safe knowledge of axioms. In Leon Horsten and Philip Welch, editors, *Gödel's Disjunction: The scope and limits of mathematical knowledge*, chapter 10, pages 243–253. Oxford University Press, Oxford and New York, 2016.

Seth Yalcin. *Modality and inquiry*. PhD thesis, MIT, 2008.

Seth Yalcin. Figure and ground in logical space. Posted at `http://escholarship.org/uc/item/11c0x4n5`, on 12–02–2015, 2015.

Seth Yalcin. Belief as question-sensitive. *Philosophy and Phenomenological Research*, 2016. ISSN 1933-1592. doi: 10.1111/phpr.12330. URL `http://dx.doi.org/10.1111/phpr.12330`.