

GRANULARITY PROBLEMS

Jens Christian Bjerring and Wolfgang Schwarz

11 April 2015

1 INTRODUCTION

If mental or linguistic content is modelled in terms of possible worlds it becomes impossible to distinguish between necessarily equivalent contents. Since ‘ $2+2=4$ ’ and ‘there are infinitely many primes’, for example, are both true at all possible worlds, they come out as having the same content. This is widely regarded as an embarrassment. After all, the two sentences seem to say very different things; they are not cognitively equivalent; they have different communicative effects; they are not interchangeable in attitude reports, subjunctive conditionals, and other embedded contexts.

An attractive strategy to overcome these drawbacks is to extend logical space by “impossible possible worlds”, making room for more fine-grained, *hyperintensional* possible-worlds propositions. If there are worlds where $2+2$ does not equal 4 and others where there are only finitely many primes, the truth-values of ‘ $2+2=4$ ’ and ‘there are infinitely many primes’ no longer coincide throughout the extended space of worlds. The two sentences then express different possible-worlds propositions.

The proposal is not just a technical fix. It also captures the intuition that for ordinary mortals like us, the space of possibilities is larger than classical logical space. We do not know all logical consequences of what we know. Perhaps Goldbach’s Conjecture is entailed by the Peano axioms, but we still do not know it. Both the conjecture and its negation are live possibilities for us and both of them can be represented in our language and thought. Yet, in classical logical space, all possible worlds make Goldbach’s Conjecture true, if it is true, or they all make it false, if it is false. By including in logical space impossible worlds where the conjecture is true and others where it is false, we seem to account for possibilities that, for us, remain live candidates for actuality.¹

On closer inspection, it turns out that tracking hyperintensional distinctions with impossible worlds is not as straightforward as it may at first appear. In sections 2 and 3 we argue that the classical possible-worlds framework is an essentially coarse-grained framework. If this is true, it means that we cannot deal with hyperintensionality by somehow extending or augmenting the classical framework with impossible worlds; rather,

¹ Nolan [2014] argues that impossible worlds may have further uses to illuminate metaphysical features of the world such as essence, grounding, and properties.

we need a completely new framework. But in pursuit of such a framework, as we argue in section 4, we should bear in mind that overly fine-grained content is as undesirable as overly coarse-grained content. For the theoretical and pre-theoretical roles associated with mental and linguistic content prohibit an excessively fine-grained individuation, which, for instance, would see a difference in content for any difference in morphology. An adequate notion of content, we claim, should have *intermediate granularity*. In section 5, however, we argue that there is no such notion of content: any way of assigning content to linguistic or mental items must be either too coarse-grained or too fine-grained.

2 IMPOSSIBLE-WORLDS SEMANTICS

Let’s begin with the idea that fine-grained differences in content could be captured in the possible-worlds framework by extending the traditional space of possible worlds, allowing for worlds where $2+2$ is 5 or where the number of primes is finite. A well-known challenge for this strategy is how to construct the relevant impossible worlds. Here we will set this issue aside, although we consider it to be an open problem.²

For concreteness, we will occasionally refer to a very simple construction on which possible and impossible worlds are identified with sets of sentences. A sentence S then counts as *true at* a world w —equivalently, w *verifies* S —iff S is a member of w . The account has obvious difficulties handling ambiguity, vagueness and context-dependence (among other things), but it can serve as a simple illustration for cases where these phenomena can be set aside.³

It appears to be a common sentiment among friends of impossible worlds that an extension of logical space with impossible worlds constitutes a fairly conservative or moderate extension of the classical framework. According to Jago, for instance, impossible worlds allow us to incorporate hyperintensionality into the possible-worlds framework “while preserving its best features”. ([Jago 2014: 14].) We disagree. Once impossible

² See [Jago 2014] for a detailed discussion. Jago suggests that possible and impossible worlds can be identified with arbitrary sets of sentences in a restricted Lagadonian language \mathcal{L} . An English sentence S is then true at a world w iff w contains an \mathcal{L} -translation of S . Problems arise in cases where different English sentences translate into the same \mathcal{L} -sentence. Arguably, ‘woodchucks are whistle-pigs’ and ‘woodchucks are woodchucks’ have the same Lagadonian translation, as do ‘ $\text{CLXVII} = 167$ ’ and ‘ $167 = 167$ ’. But, as Jago himself points out elsewhere, there are contexts in which we want to treat these sentences as having different contents.

³ In the literature on impossible worlds, there is a standard distinction between “American” and “Australian” type impossible worlds (cf. [Berto 2013]). Roughly speaking, truth at an Australian world is closed under some non-classical consequence relation, while American worlds are not subject to any such restriction. For our purposes, only American constructions will be relevant. To model the kinds of possibilities that seem to be live possibilities for ordinary people, we need worlds where A is true but B is false, even when A logically entails B in some non-classical logic. (Unless the non-classical entailment relation is so weak to blur the distinction between Australian and American worlds.)

worlds are included in logical space, we do lose the best features of the possible-worlds framework. In particular, we lose much of the appeal of the framework in compositional semantics, and we have to give up the analysis of belief and information in terms of exclusion of possibilities.⁴

We begin with compositional semantics. Standard possible-worlds semantics has a simple story to tell about the Boolean connectives: the possible-worlds content of a conjunction ‘ A and B ’ is the intersection of the contents of A and B ; disjunction and negation correspond similarly to set-theoretic union and complementation. What becomes of this story in our extended logical space? Arguably, it must be given up. For the set-theoretic rules just stated imply that any sentences that are equivalent in standard propositional logic are true in exactly the same possible worlds. So all propositional tautologies come out as having the same possible-worlds content. But this is just what we want to avoid on the hyperintensional picture: it should not be the case that all propositional tautologies, no matter how complex, have the same content.

So we can no longer give the semantics for conjunction, disjunction, and negation in terms of the familiar set-theoretic operations. In the extended space of worlds, there will, for example, be worlds verifying both A and B but not their conjunction ‘ A and B ’. And so the possible-worlds proposition associated with ‘ A and B ’ is not the intersection of the propositions associated with A and B .

Something similar happens in the case of modals. The semantics of ordinary-language modals is rather complex, so let us focus on a philosopher’s dialect of English, in which ‘Necessarily A ’ attributes metaphysical necessity to the proposition expressed by the embedded sentence A . If all worlds are metaphysically possible, we can simply say that ‘Necessarily A ’ is true at a world w iff A is true at all worlds. Now when we go hyperintensional, we want to assign different contents to sentences such as ‘Necessarily, $2+2=4$ ’ and ‘Necessarily, there are infinitely many primes’. After all, these sentences seem to say very different things, they are not interchangeable in attitude reports, etc. So we need impossible worlds where one of them is true and the other false. Since ‘ $2+2=4$ ’ and ‘there are infinitely many primes’ are true at all metaphysically possible worlds, it follows that we can no longer say that ‘Necessarily A ’ is true at an arbitrary world w iff A is true at all metaphysically possible worlds. Worse, the truth-value of ‘Necessarily A ’ is no longer determined in any systematic way at all by the set of worlds at which A is true. To see why, let A be a reasonably complex metaphysical necessity, true at some impossible worlds and false at others. Let w be some impossible world that verifies A . Does w verify ‘Necessarily A ’? We cannot tell: some impossible A -worlds also verify ‘Necessarily A ’, others do not. In sum, we can no longer give an informative account of the conditions under which ‘Necessarily A ’ is true at a world in terms of the proposition

⁴ We have no objections to the “Australian” use of impossible worlds in the model theory of non-classical logic (see e.g. [Restall 1997]).

expressed by A .

Parallel arguments can be made for counterfactuals, attitude reports, and other constructions where possible worlds have proved useful in compositional semantics. In each case, the (putatively) well-behaved compositional behavior of possible-worlds propositions is lost when the space of worlds is extended to include impossible worlds.⁵

On reflection, these observations should come as no surprise. (They will hardly be news for advocates of impossible worlds.) Standard rules of compositional semantics reduce the possible assignments of truth-values to the sentences in a language. They entail that if A is true, then ‘ A or B ’ cannot be false; that if ‘someone fears everyone’ is true, then ‘no one fears themselves’ must be false; that if the Peano axioms are true, then Fermat’s Last Theorem (presumably) cannot be false. In our extended space of worlds, we do not want these entailments. We want to allow for worlds where things are true although their consequences are false.

So once we extend the space of possible worlds, we cannot hold on to the compositional rules of classical possible-worlds semantics—however elegant and powerful these rules may be in the compositional analysis of meaning. Of course, the classical rules remain valid in the restricted set of metaphysically possible worlds, but for the whole space, we need to come up with alternatives. And here we worry that the alternatives will lack the appeal of the classical rules. For instance, in our toy construction of worlds as sets of sentences, the meaning of ‘and’ maps any pair of sets $\{w : A \in w\}$ and $\{w : B \in w\}$ to the corresponding set $\{w : ‘A \text{ and } B’ \in w\}$. This operation looks nothing like familiar set intersection, and there is something uninformative or trivial about the new rule. Intuitively, it should make a big difference whether ‘and’ expresses conjunction or disjunction, yet the rule

⁵ The case of counterfactuals is noteworthy because it is often taken to *support* a hyperintensional construction of logical space, on the grounds that this allows assigning different truth-values to counterfactuals with impossible antecedents. The basic idea is to say, in line with classical possible-worlds accounts, that

(*) $A \Box \rightarrow C$ is true at w iff the closest A -worlds to w are C -worlds.

By including impossible worlds among the closest A -worlds, we can allow that ‘if I had squared the circle, I would be famous’ might be true while ‘if I had squared the circle, nobody would care’ is false (see e.g. [Nolan 1997]). At first glance, this looks like a nice, compositional story. However, for which worlds w is (*) meant to hold? If ‘ w ’ ranges only over metaphysically possible worlds, we still need to be told under what conditions a counterfactual is true at an impossible world. (Given that counterfactuals can embed other counterfactuals, (*) would even fail to determine the truth-conditions of counterfactuals at metaphysically possible worlds.) However, if ‘ w ’ ranges over all worlds, then (*) is problematic for reasons parallel to those we discussed for ‘Necessarily’. The problem is especially vivid if we assume—as seems plausible—that (*) validates basic principles of conditional logic such as modus ponens and modus tollens. For then it will follow that truth at impossible worlds is closed under these principles. So we could not leave room for agents, who, for example, believe A and $A \Box \rightarrow (C \Box \rightarrow \neg A)$, without believing $\neg C$, since every world that verifies the first two sentences would verify $\neg C$.

just stated is correct either way!

3 IMPOSSIBLE-WORLDS PRAGMATICS

Arguably, the main appeal of the possible-worlds framework lies not so much in the resources it provides for compositional semantics, but in its promise of providing a unified, systematic account covering mental, informational, and linguistic content, and various interactions between these notions.

One starting point here is the idea that to acquire information is to exclude possibilities. When we learn that Bob is in Rome, we can exclude the possibility that he is in Paris. Possibilities—whatever they are—can be ordered by specificity: that Bob is in Rome is a more specific possibility than that he is in Italy; that he is in Rome on a business trip is more specific than that he is in Rome. Under plausible (although non-trivial) assumptions about the specificity ordering, every possibility uniquely corresponds to a set of maximally specific possibilities.⁶ So we can identify the possibility that Bob is in Rome with a set of maximally specific possibilities, namely those which entail that Bob is in Rome. A maximally specific possibility is a complete way things might be—a possible world. The information we receive can therefore be modelled as the set of possible worlds “at which” Bob is in Rome, a set that excludes (in the set-theoretic sense) all possibilities at which Bob is somewhere else. Similarly, the totality of an agent’s knowledge can be modelled as a set of possible worlds, comprising all possibilities that might, for all the agent knows, be actual ([Hintikka 1962]).

In the same spirit, Stalnaker [1970] proposed a possible-worlds model for the dynamics of assertion. Before Bob told Alice that he is in Rome, the contextually open possibilities did not settle Bob’s location; perhaps they included worlds where Bob was in Paris and others where he was in Rome. Bob’s utterance of ‘I am in Rome’ then had the effect that all worlds where he is not in Rome got removed from the set of contextually open possibilities. Had Bob uttered a different sentence—say, ‘I am in Paris’—the set of contextually open possibilities would have changed in a different, but equally predictable manner. The sentences thus have different “context change potential”.

These ideas have been successfully extended and refined to analyze a large variety of phenomena. However, as they stand, they are insensitive to hyperintensional distinctions. The possible-worlds model of information and knowledge can’t account for agents who know simple logical truths without knowing all logical truths. Stalnaker’s model of

⁶ The required assumptions are the conditions on a complete, atomic Boolean lattice. The isomorphism between possibilities and sets of maximally specific possibilities is an instance of Stone’s representation theorem—bracketing a technicality concerning non-principal ultrafilters that arises if the space of possibilities is infinite (see e.g. [Cresswell 2006]).

assertion seemingly can't explain the different communicative effects of sentences that are true at the very same worlds. It is here that impossible worlds promise relief.

Yet, as soon as we try to adapt the above picture to an impossible-worlds framework, we run into severe difficulties. [Bjerring 2013] points out the following. Assume worlds—whether possible or impossible—are *maximal* in the sense that for every sentence they verify either it or its negation. Consider a world w that verifies some complex contradiction C of, say, classical propositional logic. Since C is a contradiction, there is a proof of $\neg C$. That is, there is a sequence of sentences S_1, \dots, S_n , ending in $S_n = \neg C$, each member of which is either a simple tautology or derivable from one or two earlier elements in the sequence by a simple logical rule like modus ponens. Given that worlds are maximal, w contains either S_i or $\neg S_i$ for each element in the sequence S_1, \dots, S_n . So there are exactly three possibilities for w : either (i) w verifies the negation of some simple tautology, or (ii) w verifies the premises of a simple logical rule as well as the negated conclusion, or (iii) w verifies both C and $\neg C$. In each case, w is a trivially inconsistent world by the standards of classical propositional logic.

So we cannot allow for situations in which all trivially inconsistent possibilities have been ruled out while some non-trivially inconsistent possibilities remain open. Nor can we model the belief states of logically non-omniscient, yet moderately competent agents who can rule out all trivially inconsistent possibilities without ruling out all non-trivially inconsistent possibilities. To avoid logical omniscience, some inconsistent possibilities must remain live possibilities for such agents, but then some trivially inconsistent possibilities must too.

One might try to circumvent this problem by allowing for non-maximal (“partial”) worlds that verify neither S nor $\neg S$, for some sentence S .⁷ There can then be worlds that verify the simple tautologies at the beginning of S_1, \dots, S_n but that verify neither S_i nor $\neg S_i$ for more complex tautologies S_i further down the sequence. But recall that possible worlds are *maximally specific* ways things might be. How can worlds that verify neither S nor $\neg S$ be maximally specific ways things might be, by the lights of a moderately rational agent? Well, the agent might (rationally or irrationally) believe that S and $\neg S$ are both false. In that case, worlds in which either S or $\neg S$ is true are plausibly incompatible with the agent’s beliefs. Similarly if the agent merely reserves some credence for the possibility that S and $\neg S$ are both false: her space of doxastically possible worlds should then include partial worlds that verify neither S nor $\neg S$.

The problem is that failure of logical omniscience can hardly be reduced to skepticism about bivalence. Consider an agent who is certain that either S_i or $\neg S_i$ is true, for all members of the sequence S_1, \dots, S_n . Worlds that verify neither S_i nor $\neg S_i$ should then not count as live possibilities for her: they are not maximally specific ways things might be. But even in that context, the complex contradiction C may be deemed possible by

⁷ This is the response taken in, for instance, [Jago 2014].

the agent, and its negation $\neg C$ may provide her with non-trivial information. This time, the worlds she rules out cannot be partial worlds, since those were already ruled out from the start. We are left with the original problem.⁸

The upshot is that we cannot use impossible worlds to characterize the knowledge or belief of logically competent but non-omniscient agents—not if we want to understand these worlds as maximally specific ways things might be. Either the worlds are blatantly inconsistent, in which case they don’t represent genuine possibilities (by the lights of the agent), or they are partial, in which case they don’t represent maximally specific ways things might be, assuming the agent accepts bivalence.

There is a more general point here. Recall that in the classical possible-worlds framework, the identification of possibilities (or propositions) with sets of possible worlds was justified by certain structural assumptions about the space of possibilities, namely that the possibilities ordered by specificity constitute a complete, atomic Boolean lattice. If we move to hyperintensional possibilities, these structural assumptions become highly implausible. In fact, it is not entirely clear what should now count as the specificity order. Is the possibility expressed by ‘Bob is in Paris on a business trip’ still more specific than that of ‘Bob is in Paris’? We could say that it is, on the grounds that the former entails the latter. But on this approach, the anti-symmetry condition on the specificity order fails: there will be distinct hyperintensional possibilities that entail one another. As a consequence, the possibilities will not correspond to sets of maximally specific possibilities (see [Pollard 2008], [Pollard 2011]). Alternatively, we could try to construe the specificity order in some more fine-grained, hyperintensional manner. For instance, we could say that A counts as more specific than B only if it is *obvious* that B follows from A . But then we lose even more of the structural conditions on a Boolean lattice. (The “obvious entailment” relation is not even transitive.)

Popular constructions of impossible worlds tend to obscure these facts. In our naive construction of worlds as sets of sentences, there still seems to be a natural correspondence between the hyperintensional possibilities expressed by sentences and sets of worlds, mapping every sentence A to the set of worlds that contain A . However, in sharp contrast to the traditional possible-worlds framework, the entire set-theoretic structure here does no work. The *expressible* propositions—those that correspond to sentences—are always set-theoretically independent: the worlds that verify A never form a subset of the worlds that verify another sentence B . Inexpressible propositions which do stand in non-trivial set-theoretic relations to expressible ones seem to be mere artifacts of the construction.

⁸ The problem is not that we might have to accept what Jago [2014] calls “epistemic oversights” ([Jago 2014: 206]): that logically competent but non-omniscient agents fail to believe even trivial consequences of things they believe. The problem is how we are supposed to model epistemic oversights with possible worlds. According to Jago, the agents’ doxastic space “has to contain worlds that are highly incomplete” (ibid.). But then the worlds can’t be understood as maximally specific ways things might be by the lights of the agent.

Consider, for instance, the set of worlds that verify ‘ $2+2=4$ ’ conjoined with some other worlds that verify ‘ $2+2=5$ ’. Is this supposed to be a possibility? Is it a possible object of belief? Do you automatically believe it whenever you believe that $2+2=4$?

What is really modelled by the hyperintensional construction is an account on which different contents are simply independent entities, and where learning something amounts to *adding it* to a stock of previously learned contents. In our simple construction, when we model an agent’s belief state by a set of worlds, the only aspect of the model with real significance is which sentences are verified by all worlds in the set. These are the sentences the agent believes. When the agent learns another sentence, the set of sentences verified by all worlds grows by one. Superficially, learning A is still modelled as excluding worlds that do not verify A . But these worlds are not complete ways things could be. They are not possibilities at all. The exclusion operation on the space of worlds is just a roundabout way of representing the addition of a new sentence.

All this need not show that the hyperintensional account does not work. It only shows that the account should not be regarded as a moderate extension of the traditional possible-worlds account. It is a completely different approach, disguised as a moderate extension.

Now one might argue that a radically different approach is indeed required to account for the hyperintensionality of mental and linguistic content. But before we jump to a particular alternative, let us try to get clear about the goal. If possible-worlds propositions are too coarse-grained, what grain size would be adequate?

4 THE NEED FOR COARSE-GRAINING

Our toy construction of worlds as sets of sentences makes possible-world propositions extremely fine-grained. No two sentences ever express the same content, since there are always worlds that contain one but not the other. However, an excessively fine-grained individuation of content is as problematic as an excessively coarse-grained one.

For one thing, intuitively we do think that at least some sentences have the same content. ‘ $2+2=4$ ’ and ‘there are infinitely many primes’ say quite different things, but that is not true for pairs such as ‘I nearly fell’ and ‘I almost fell’, for ‘ $3 < 9$ ’ and ‘ $9 > 3$ ’, or for ‘it is raining’ in English and ‘il pleut’ in French. In these pairs, both sentences intuitively express the very same thought and make the same claim about reality.

Second and more seriously, consider Euclid’s discovery that there are infinitely many primes. Presumably the content of Euclid’s discovery can be expressed not only by ‘there are infinitely many primes’ but also by trivially equivalent statements such as ‘the number of primes is infinite’. That the number of primes is infinite was not a further discovery, also made by Euclid. Similarly, when the Babylonians discovered that Hesperus is Phosphorus, what they discovered could just as well be expressed by ‘Phosphorus is

Hesperus’ (but not by ‘Hesperus is Hesperus’). In either case, the content of the discovery seems to have intermediate granularity.

The same is true for other attitudes. For example, you cannot notice that I nearly fell without noticing that I almost fell: these statements attribute the very same state of noticing. So if noticing is a relation between a subject and a content, then the relevant content is not as fine-grained as the words we can use to express it.

Third, overly fine-grained conceptions of content would preclude sentences in different languages from having the same content. If different sentences never agree in content, it would be impossible to express in English what one can say in French by uttering ‘il pleut’. If ‘*A* believes that *S*’ attributes a belief whose content is that of the embedded sentence *S*, ‘Euclid believed that there are infinitely many primes’ would attribute to Euclid a belief that can only be expressed or attributed to him in English—which would raise the question how Euclid could have acquired such an “essentially English” belief, given that neither he nor anyone else at his time spoke English.

That last remark leads to a fourth argument against excessively fine-grained conceptions of content: such conceptions make it mysterious how mental and linguistic types get to have their content. Suppose the content of a mental state is determined by causal relations to the environment, behavioural dispositions, inferential links, and further features along these lines. It is then hard to see how there could be a genuine difference in content for any two sentences that might be used to attribute an attitude. For example, what facts about Euclid’s cognitive state could possibly have made it true that he believed that there are woodchucks without believing that there are whistle-pigs?⁹ And if mental content is relatively coarse-grained, then many foundational accounts of linguistic content—for example, in the tradition of [Grice 1957]—imply that linguistic content will be equally coarse-grained.

Finally, intermediately-grained notions of content are central to many projects in philosophy. For example, a key idea in Frege’s *Grundlagen* [1884] is that quantified mathematical statements sometimes agree in content with non-quantified statements, as in the case of (1a) and (1b):

- (1) a. *A* and *B* are parallel.
- b. *A* and *B* have the same direction.

That would be trivially false if syntactically different sentences could never agree in content.

One problem that emerges from all these considerations is that there is unlikely to be a single level of granularity that works for every purpose, in every context. The kind of content established by a formal proof in, say, intuitionistic logic, is a lot more fine-grained

⁹Stalnaker 1976 even argues that functionalist accounts of mental content only deliver the very coarse-grained contents of traditional possible-worlds accounts.

than the content seemingly expressed by ordinary assertions or the content represented by a map. When we talk about beliefs, we often want to treat the propositions that there are woodchucks and that there are whistle-pigs as identical, especially when the subject of the belief does not speak English. But not always (see [Ripley 2012]). This suggests that there is no way of assigning contents to sentences that will get all cases right.¹⁰

Frege put forward a criterion for identity of content that—albeit inadvertently—takes into account some such flexibility and relativity. In essence, Frege’s proposal is that two sentences have the same content if and only if one could not regard one of them as true and the other as false. Sentences that satisfy this condition Frege called *equipollent* (see e.g. [Frege 1891: 14], [Frege 1892: 47] and [Frege 1983: 152f.]).¹¹ By Frege’s criterion, ‘ $2+2=4$ ’ and ‘there are infinitely many primes’ plausibly come out as having different content, while ‘ $3 < 9$ ’ and ‘ $9 > 3$ ’, or (1a) and (1b) have the same content.

In general, however, whether two sentences count as equipollent will depend on what kinds of agents we consider when we ask whether it is possible to believe that the sentences might have different truth-values. The individuation of content becomes relative to a base level of information and cognitive capacities. Consider (1a) and (1b). Most speakers of English would probably dismiss the hypothesis that one of these might be true while the other is false, but it is not hard to imagine speakers for whom this is a live possibility. On a somewhat more advanced level, consider the equations ‘ $x = \ln y$ ’ and ‘ $y = e^x$ ’. While an algebra student may wonder whether they can differ in truth-value, her teacher may find it utterly obvious that they cannot. In the limit, if we only consider ideal agents with unbounded cognitive capacities, equipollence might reduce to something like necessary or a priori equivalence. At the other extreme, for utterly confused agents—or, less interestingly, for agents who do not speak the relevant language—it reduces to a trivial relation that never holds between different sentences.

We can therefore see Frege’s proposal as offering a whole range of criteria whose end-points are the extremely coarse-grained individuation of classical intensional semantics and the extremely fine-grained individuation of our hyperintensional alternative. Since it is doubtful that a single grain size is adequate in all contexts and for all purposes, some such pluralism about content might be just what we need.

Now Frege’s criterion only tells us when two sentences have the same content. It does not tell us what these contents are. Unfortunately, we will see that it is impossible to construe a notion of content that satisfies Frege’s criterion for intermediary points along the scale, where we consider the judgements of moderately competent, but logically

¹⁰ Stalnaker raises versions of the present worry in [Stalnaker 1991] and [Stalnaker 1999b].

¹¹ We will not enter into the exegetical details about what exactly Frege meant by equipollence, and how his proposal squares with other remarks in which he seems to suggest that the sense of a sentence is composed of the senses of its parts; see e.g. [Penco 2003], [Kemmerling 2010], and [Schellenberg 2012] for discussion.

non-omniscient agents.

5 THE INTRANSITIVITY OF SAMENESS OF CONTENT

Consider another lengthy sequence S_1, \dots, S_n of sentences. This time assume that each S_{i+1} is trivially equivalent to its predecessor S_i —meaning that each trivially follows from the other—although S_n is not trivially equivalent to S_1 . For example, S_1, \dots, S_n might be a sequence of algebraic equations, where each step is a trivial transformation of the previous equation, although S_n is a highly non-trivial transformation of S_1 . For moderately competent agents, subsequent equations should plausibly count as equipollent: the hypothesis that one is true and the other false can be immediately ruled out. Yet, it may remain a live possibility for such agents that the first equation is true but the last one false. So S_1 and S_n are not equipollent. By Frege’s criterion, it now follows that each sentence in the sequence has the same content as its immediate neighbors, although the first and the last sentence do not. Evidently, this is impossible.

Note that the problem here does not rely on any assumptions about the nature of contents—whether they are sets of worlds, structured entities, mental states, or *sui generis* whatnots. The problem is that equipollence is intransitive, while identity is transitive. So it cannot be true, as Frege’s criterion requires, that two sentences have identical content just in case they are equipollent. The problem disappears only at the end points of the Fregean spectrum. For ideal agents who instantly and effortlessly recognize every consequence of every sentence, equipollence may well be transitive. The same is true for utterly confused agents for whom no sentence is equipollent to any other.

It would be short-sighted to blame Frege’s criterion. The general problem does not turn on Frege’s particular individuation of intermediately-grained content. Return to our sequence S_1, \dots, S_n . Any assignment of content must cut the sequence into equivalence classes of sentences with the same content. If we are looking for an intermediately-grained notion of content, we do not want too many cuts in the sequence. In particular, we do not want to say that no two sentences in the sequence have the same content. We also do not want to say that they all have the same content. But then it gets hard to justify the cuts. Suppose the first cut is after S_{10} . So S_1 and S_{10} count as having the same content, while S_{10} and S_{11} count as having different contents, despite the fact that S_{10} and S_{11} are more obviously equivalent than S_1 and S_{10} .¹²

¹² The problem is not vagueness. It is tempting to say that (for example) S_1 and S_2 determinately agree in content, that S_1 and S_{10} are a borderline case, and that S_1 and S_{20} determinately do not agree in content. There is nothing especially troublesome about a vague notion of content. But in order for there to be even a vague assignment of content to sentences, the precisifications of the same-content relation must be equivalence relations—and that’s what we don’t have. By analogy, suppose someone proposes a coarse-grained notion of colours by the principle that x and y have the same coarse-grained colour iff the shade of x is similar to that of y . All precisifications of the vague right-hand side are

The same line of argument could be made with sentences S_1, \dots, S_n across different languages, where each pair of adjacent sentences seem to agree in meaning, although S_1 and S_n do not. For a simple example, consider a context in which we want to distinguish between the proposition that there are woodchucks and the proposition that there are whistle-pigs (following [Ripley 2012]). In German there is only one word for woodchucks: ‘Waldmurmeltier’. So both ‘there are woodchucks’ and ‘there are whistle-pigs’ translate into ‘es gibt Waldmurmeltiere’. Now which of these three sentences have the same content?

For another variation, consider a sequence of belief reports R_1, \dots, R_n —in different languages, perhaps—where in each adjacent pair the complement sentences that specify the believed content are trivially equivalent, although the complement of R_1 is not trivially equivalent to that of R_n . Again, we face the same uncomfortable choice between saying that practically any change in complement sentence attributes a different belief, or making isolated cuts where a small change in complement sentence amounts to a different belief, even though other, intuitively larger changes do not.

Of course, there are systematic ways of placing the cuts. The classical neo-Russellian account, for example, sees a difference in content whenever there is either a difference in syntactic structure or a difference in reference. Consequently, ‘Hesperus is Hesperus’ and ‘Hesperus is Phosphorus’ are assigned the same content, while ‘ $3 < 9$ ’ and ‘ $9 > 3$ ’, or ‘it is raining’ in English and its Russian translation ‘Идёт дождь’ (literally, ‘goes rain’) are assigned different contents. The problem is that this way of individuating content does not fit any of the phenomena surveyed in the previous section that seemed to call for an intermediately-grained notion of content.

The same is true for David Chalmers’s recent proposal in [Chalmers 2011] and [Chalmers 2012: 248ff.], on which two sentences have different contents iff they either differ in syntactic structure or some of their constituents differ in (primary) intension or extension. ‘Hesperus is Phosphorus’ and ‘Hesperus is Hesperus’ now plausibly come out as having different contents. However, ‘ $3 < 9$ ’ and ‘ $9 > 3$ ’, ‘it is raining’ and ‘Идёт дождь’, or Frege’s (1a) and (1b) still have different contents. *Pace* Chalmers, his notion of content is therefore not very Fregean. Looking back at the phenomena in the previous section, Chalmers’ notion of content almost always cuts too finely. Sometimes it cuts too coarsely. For example, if we define ‘ ν ’ to denote the smallest positive number x for which $\cos(x/2) = 0$, then ‘ $\nu = \pi$ ’ and ‘ $\pi = \pi$ ’ have the same Chalmersian content, although the former seems informative but the latter trivial.¹³

intransitive, and so none of them determines a partition of objects by coarse-grained colour.

¹³ Chalmers’s individuation of content resembles Carnap’s individuation in terms of intensional isomorphism. The present objection to Chalmers is raised in [Church 1954] as an objection to Carnap. Chalmers mentions the problem in [Chalmers 2012: 249] and replies that “it is at least arguable that [ν] should be understood to have complex structured content”, which would distinguish it from ‘ π ’. But then the relevant structure cannot be tied to syntax or logical form, which is what Chalmers’s

The arguments we have outlined can be applied to any attempt to assign contents to sentences. We can always construct sequences of sentences in which—according to the intuitive and theoretical considerations we have reviewed—all adjacent sentences should have the same content while sentences that are sufficiently far apart should have different content. Since it is impossible to meet both these requirements, no assignment of content can get all the data right.

6 CONCLUSIONS

We have tried to establish two main claims. First, the framework of possible worlds is essentially coarse-grained: if we want to distinguish logically equivalent contents, nothing but confusion—and a superficial adherence to tradition—is gained by modeling these contents as sets of worlds. Second, while a variety of phenomena seems to call for a notion of content more fine-grained than sets of possible worlds, and more coarse-grained than linguistic morphology, the required intermediately-grained notion of content does not exist. In particular, no assignment of content to sentences can satisfy Frege’s equipollence criterion or our pre-theoretic criteria for sameness of content. Thus it is no surprise that popular accounts of semantic and epistemic content tend to be either implausibly coarse-grained or implausibly fine-grained (or both).

This leaves two options for semantic theorizing. One is to drop the assumption that semantic facts can be captured by assigning to linguistic or mental items some kind of extra-linguistic content. If instead we confine ourselves to studying relations within the domain of the linguistic or the mental—relations of same-saying, synonymy or equipollence, for example—then it does not matter whether these relations are transitive.

Alternatively, we can continue assigning content to linguistic and mental items, accept that our assignment is implausibly coarse-grained or implausibly fine-grained (or both), and try to give a more complex analysis of the phenomena that seem to call for an intermediately-grained notion of content. In particular, we could give up the assumption that attitude reports and speech act reports simply state a relation between the subject and the proposition expressed by the embedded sentence. The hypothesis that ‘Hesperus is Phosphorus’ and ‘Hesperus is Hesperus’ express the same proposition then no longer makes the false prediction that the two sentences are interchangeable in attitude reports: it no longer follows that anyone who believes that (or asserts that) Hesperus is Hesperus also believes that (or asserts that) Hesperus is Phosphorus.

The present idea has been defended by philosophers such as Salmon [1986] or Stalnaker [1991] who advocate a coarse-grained account of content. The considerations of the

official proposal assumes. If even syntactically simple terms can have structure, one would like to know a lot more about how that structure is determined. Does ‘ π ’ have structured content? Which of the many equivalent definitions of π is reflected in its structure?

present paper suggest that it is unavoidable. No conception of content—no matter how coarse-grained or fine-grained—can fit the identity conditions apparently imposed by our ordinary judgements about attitudes and speech acts.

REFERENCES

- Francesco Berto [2013]: “Impossible Worlds”. In Edward N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy*, Winter 2013 edition
- Jens Christian Bjerring [2013]: “Impossible worlds and logical omniscience: an impossibility result”. *Synthese*, 190(13): 2505–2524
- David Chalmers [2011]: “Propositions and Attitude Ascriptions: A Fregean Account”. *Nous*, 45: 595–639
- [2012]: *Constructing the World*. Oxford: Oxford University Press
- Alonzo Church [1954]: “Intensional Isomorphism and Identity of Belief”. *Philosophical Studies*, 5(5): 65–73
- Max Cresswell [2006]: “From Modal Discourse to Possible Worlds”. *Studia Logica*, 82(3): 307–327
- Gottlob Frege [1884]: *Die Grundlagen der Arithmetik*. Beslau: M.& H. Marcus
- [1891]: *Funktion und Begriff*. Jena: Hermann Pohle
- [1892]: “Über Sinn und Bedeutung”. *Zeitschrift für Philosophie und philosophische Kritik*, 100: 25–50
- [1983]: *Nachgelassene Schriften*. Hamburg: Meiner, 2nd edition
- Paul Grice [1957]: “Meaning”. *The Philosophical Review*, 66(3): 377–388
- Jaakko Hintikka [1962]: *Knowledge and Belief: An introduction to the logic of the two notions*. Ithaca: Cornell University Press
- Mark Jago [2014]: *The Impossible*. Oxford: Oxford University Press
- Andreas Kemmerling [2010]: “Thoughts without parts: Frege’s doctrine”. *Grazer Philosophische Studien*, 82: 165–188
- Daniel Nolan [1997]: “Impossible Worlds: A Modest Approach”. *Notre Dame Journal of Formal Logic*, 38: 535–572
- [2014]: “Hyperintensional Metaphysics”. Forthcoming in *Philosophical Studies*

- Carlo Penco [2003]: “Frege: two theses, two senses”. *History and Philosophy of logic*, 24(2): 87–109
- Carl Pollard [2008]: “Hyperintensions”. *Journal of Logic and Computation*, 18(2): 257–282
- [2011]: “Are (Linguists’) Propositions (Topos) Propositions?” In *Logical Aspects of Computational Linguistics*, Springer, 205–218
- Greg Restall [1997]: “Ways Things Can’t Be”. *Notre Dame Journal of Formal Logic*, 38: 583–597
- David Ripley [2012]: “Structures and circumstances: Two ways to fine-grain propositions”. *Synthese*, 189(1): 97–118
- Nathan Salmon [1986]: *Frege’s Puzzle*. Cambridge (Mass.): MIT Press
- Susanna Schellenberg [2012]: “Sameness of Fregean sense”. *Synthese*, 189(1): 163–175
- Robert Stalnaker [1970]: “Pragmatics”. *Synthese*, 22. In [Stalnaker 1999a]
- [1976]: “Propositions”. In A.F. MacKay and D.D. Merrill (Eds.) *Issues in the Philosophy of Language*, New Haven: Yale University Press, 79–91
- [1991]: “The Problem of Logical Omniscience I”. *Synthese*, 89
- [1999a]: *Context and Content*. Oxford: Oxford University Press
- [1999b]: “The Problem of Logical Omniscience II”. In [Stalnaker 1999a], 255–273