

Hyperintensional semantics: a Fregean approach

Jens Christian Bjerring · Mattias Skipper

Penultimate draft, forthcoming in *Synthese*

Abstract. In this paper, we present a new semantic framework designed to capture a distinctly *cognitive* or *epistemic* notion of meaning akin to Fregean senses. Traditional Carnapian intensions are too coarse-grained for this purpose: they fail to draw semantic distinctions between sentences that, from a Fregean perspective, differ in meaning. This has led some philosophers to introduce more fine-grained *hyperintensions* that allow us to draw semantic distinctions among co-intensional sentences. But the hyperintensional strategy has a flip-side: it risks drawing semantic distinctions between sentences that, from a Fregean perspective, do *not* differ in meaning. This is what we call the ‘new problem’ of hyperintensionality to distinguish it from the ‘old problem’ that faced the intensional theory. We show that our semantic framework offers a joint solution to both these problems by virtue of satisfying a version of Frege’s so-called ‘equipollence principle’ for sense individuation. Frege’s principle, we argue, not only captures the semantic intuitions that give rise to the old and the new problem of hyperintensionality, but also points the way to an independently motivated solution to both problems.

Keywords: Hyperintensionality · Semantic content · Possible worlds · Impossible worlds · Fregean content

1 Two Problems of Hyperintensionality

When do two linguistic expressions differ in meaning, and when not? Answers to this question, arguably, depend on which aspect of meaning one has in mind. But if Frege is right, there is at least one aspect of meaning—what he called ‘sense’ (Frege 1892)—that is intimately related to the cognitive significance of linguistic expressions. Thus, when seen through a Fregean lens, the opening question becomes distinctly *cognitive* or *epistemic* in nature: whether two linguistic expressions differ in meaning comes down to whether they play suitably different roles in our cognitive lives.

Of course, Fregean notions like ‘sense’ and ‘cognitive significance’ are notoriously vague and imprecise. But Carnap (1947) proposed a way of substantiating the Fregean vision by understanding semantic content in *modal* terms. His idea was to identify the semantic content of a linguistic expression with its *intension*, defined as a function from possible worlds to extensions. Accordingly, on the Carnapian view, two linguistic expressions differ in meaning if and only if they have different intensions, that is, if and only if they are not co-extensional across all possible worlds. For example, “Mount Everest” and “the world’s highest mountain” differ in intensional content, because there *could’ve* been a mountain higher than Mount Everest, even if there *isn’t* one.

But for the purpose of capturing the kind of cognitive or epistemic aspect of meaning that Frege alluded to with his notion of sense, the intensional individuation of semantic content is too coarse-grained. Consider, for instance, the following pairs of sentences:

- (1) a. Hesperus is Hesperus.
b. Hesperus is Phosphorus.
- (2) a. $7 = 7$.
b. $(5^2 \times 211 - 4) \div 753 = 7$.
- (3) a. A .
b. $(A \wedge (\neg(\neg A \vee B))) \vee (A \wedge B)$.¹

¹Throughout we shall understand capital letters (A, B, \dots) as placeholders for inter-

These sentence pairs clearly seem to differ in cognitive significance. It can presumably be cognitively significant to learn that (2a) and (2b) are equivalent: it requires non-trivial cognitive work to establish that $(5^2 \times 211 - 4) \div 753 = 7$ is equivalent to $7 = 7$, and in establishing the equivalence, we can learn something new about mathematics. Similarly, (1a) and (1b) can presumably behave differently in conversational contexts: while uttering that ‘Hesperus is Phosphorus’ may well change the conversational presuppositions in a group of Babylonian astronomers, uttering the trivial truth that ‘Hesperus is Hesperus’ won’t. Finally, (3a) and (3b) are not substitutable *salva veritate* in propositional attitude contexts: it may well be true in a given context to say that ‘Bob believes that A ’, but false to say that ‘Bob believes that $(A \wedge (\neg(\neg A \vee B))) \vee (A \wedge B)$ ’. Thus, if we look at the roles that sentence pairs such as (1) to (3) play in our cognitive lives, there is a clear sense in which they differ in meaning.

Yet, since each pair of sentences in (1) to (3) have the same truth-value across all possible worlds, they have the same intension, and are hence associated with the same semantic content. So from a Fregean perspective, the intensional account does not make sufficiently fine-grained distinctions among contents. We will call this the ‘old problem of hyperintensionality’—old because it was recognized from the early beginnings of the intensional revolution in the second half of the twentieth century.²

While the old problem of hyperintensionality has led some philosophers to reject the worlds-based approach to semantics altogether, others have tried to solve the problem within a worlds-based framework. Perhaps the most popular strategy has been to replace Carnapian intensions with more fine-grained *hyperintensions*, defined as functions from possible or *impossible* worlds to extensions. Intuitively, while possible worlds represent ways the world *could* be, impossible worlds represent ways the world *could not* be. That is to say, impossible worlds allow for the metaphysically, analytically, mathematically,

preted sentences (‘snow is white’, ‘Joe is hungry’,...).

²See, e.g., Barwise and Perry (1983), Cresswell (1975), and Fine (2015). Already Carnap (1947) noted that attitude contexts are problematic for the intensional individuation of content and proposed to handle such contexts by appeal to a notion of ‘intensional isomorphism.’

and logically impossible to happen. For example, there might be impossible worlds at which A is true, but its logical equivalent $(A \wedge (\neg(\neg A \vee B))) \vee (A \wedge B)$ is not. This, in turn, allows us to associate (3a) and (3b) with distinct hyperintensions, and hence with distinct contents. So if we avail ourselves of impossible worlds, we seem well-equipped to solve the old problem of hyperintensionality.

However, impossible worlds can be a bit of a mixed blessing. On the one hand, if impossible worlds are to provide assistance in solving the old problem of hyperintensionality, they must allow us to draw distinctions among contents that are not licensed by possible worlds. On the other hand, if impossible worlds allow us to draw *arbitrarily* fine-grained distinctions among contents, we end up making semantic distinctions where, from a Fregean perspective, there should be none. Consider, for instance, the following pairs of sentences:³

- (4) a. The number of primes is infinite.
- b. There are infinitely many primes.
- (5) a. $x = y$.
- b. $y = x$.
- (6) a. A .
- b. $A \wedge A$.

These sentence pairs, in contrast those in (1) to (3), do *not* seem to differ in cognitive significance. It will hardly be cognitively significant to be told that (6a) and (6b) are equivalent: it requires only trivial cognitive work to establish that A is equivalent to $A \wedge A$, and in establishing the equivalence, we do not seem to learn anything new about logic. Similarly, it will hardly make a difference in ordinary conversational contexts whether one utters (4a) or (4b): uttering that ‘the number of primes is infinite’ may well have an impact on the conversational presuppositions in a group of speakers, but

³The pairs (5) and (6) are extracted from Frege (1984, pp. 390-406). While some might find the particular examples in (4) to (6) contentious, they are in our opinion well-suited to motivate the idea that Fregean content should not be arbitrarily fine-grained. See also Chalmers (2011), Jago (2014), Bjerring and Schwarz (2017), and Nolan (forthcoming) for similar examples.

a subsequent utterance of ‘there are infinitely many primes’ will presumably not have any additional impact. Finally, the sentences (5a) and (5b) are substitutable *salva veritate* in propositional attitude contexts: if it is true in a given context to say that ‘Bob believes that $x = y$ ’, it is presumably also true to say that ‘Bob believes that $y = x$ ’. Thus, if we look at the roles that sentence pairs such as (4) to (6) play in our cognitive lives, there is a clear sense in which they have the same meaning.

Yet, if there are impossible worlds at which, say, ‘ $x = y$ ’ is true but ‘ $y = x$ ’ is not, we end up associating (5a) and (5b) with distinct hyperintensions, and hence distinct contents. And the same goes for the pairs in (4) and (6). So for the purpose of capturing a notion of meaning that is intimately related to cognitive significance, the hyperintensional strategy risks individuating semantic content in an overly fine-grained manner.⁴ We call this the ‘new problem of hyperintensionality’—new because it has received comparatively little attention in the literature on semantic content individuation.⁵

The new problem of hyperintensionality can be seen as the flip-side of the old problem: while the old problem concerns overly *coarse-grained* content, the new problem concerns overly *fine-grained* content. Henceforth, we shall also say that the old problem of hyperintensionality is that of accommodating *hyperintensional* content, whereas the new problem of hyperintensionality is that of accommodating *non-trivial* content. If a semantic theory fails to associate sentence pairs like (1) to (3) with distinct contents, it fails to make room for hyperintensional content, and hence fails to solve the old problem of hyperintensionality. Conversely, if a semantic theory fails to associate sentence pairs like (4) to (6) with the same content, it fails to make room for non-trivial content, and hence fails to solve the new problem of hyperintensionality.

The aim of this paper is to develop a joint solution to the old and the new problem of hyperintensionality, and to do so within a broadly worlds-based framework. At the heart of our proposal lies a version of Frege’s so-called

⁴For further motivation of the idea that too much fine-graining is as troublesome as too much coarse-graining, see again Chalmers (2011), Jago (2014), Bjerring and Schwarz (2017), and Nolan (forthcoming).

⁵A notable exception is Jago (2014) whose proposal we shall discuss further in §6.

‘equipollence principle’ for semantic content individuation. According to this principle, two sentences A and B have the same content if and only if

they stand in such a relation that anyone who recognizes the content of A as true must thereby also recognize the content of B as true and, conversely, that anyone who accepts the content of B must immediately accept that of A . (Frege 1969/1979, pp. 197–98)

Effectively, that is, the equipollence principle individuates semantic content by appeal to an epistemic or cognitive relation of being *immediately recognizable as equivalent*. Even without a formally precise rendering of this relation, it is not hard to appreciate how it can help to solve the two problems of hyperintensionality. Intuitively, while ordinary persons like you and me can immediately recognize that pairs such as (4) to (6) are equivalent, we cannot immediately recognize that pairs such as (1) to (3) are equivalent. As such, the equipollence principle looks tailor-made to deliver the semantic distinctions that we are after.

As we shall see, a semantics built around the equipollence principle generates a whole *spectrum* of notions of content with different levels of granularity, ranging from extremely fine-grained—and potentially trivial—content to content that is individuated up to a priori equivalence. This semantic flexibility is desirable since different hyperintensional contexts are likely to call out for different levels of granularity.⁶ Moreover, by making the equipollence principle central to our semantic framework, we can circumvent a recent impossibility result by Bjerring and Schwarz (2017). Their result shows that any solution to the old and the new problem of hyperintensionality must give up the assumption that semantic content should be individuated in a *transitive* manner. And the relation of ‘being immediately recognizable as equivalent’ is, as we shall see, most naturally understood as an intransitive relation. So the equipollence principle offers an independently motivated way around Bjerring and Schwarz’s impossibility result.

Before we proceed, let us make three further remarks about the aim and scope of the paper. First, since the equipollence principle invokes a distinctly

⁶For further motivation of this point, see Bjerring and Schwarz (2017), Faroldi (2017) and Leitgeb (ms.).

epistemic constraint on content individuation, the Fregean approach to semantics differs importantly from the kinds of externalist approaches that we find in, e.g., Kripke (1980), Putnam (1975), and Burge (1979a; 1979b). Roughly, on their views, the semantic content of a linguistic expression is determined solely by the referential relations that it bears to various external or mind-independent entities in the world. Philosophers of an externalist bend sometimes complain that Fregean senses are ‘mysterious entities’ that must either be taken as primitive or cashed out in terms of equally obscure ‘modes of presentation.’⁷ In our framework, however, Fregean content will be spelled out in terms of hyperintensions that are nothing but functions from worlds to extensions. So although our aim is not to criticize the externalist approach to semantics—nor to defend the Fregean approach—the charge of obscurity should not trouble us.

Second, while we hope to show that our semantic framework can capture *one* key aspect of meaning akin to Fregean senses, we do not claim that it can do *all* the work that one might ask of a semantic theory. In particular, due to its epistemic nature, the equipollence principle might not be useful for reasoning about semantic phenomena that arguably require a non-epistemic analysis such as, say, consequence relations in natural languages. But as we see it, there is no problem here. There are many roles that one might want a semantic theory to play, and we need not assume that a single theory should be able to play all of those roles.

Finally, it bears mentioning that there are alternative stances on hyperintensionality that differ importantly from the broadly worlds-based approach that we will adopt. Perhaps most prominently, there are *structuralists* about content, who understand propositions as structured entities consisting of objects, properties, relations, and the like.⁸ Typically, structuralists agree that hyperintensionality is a problem, and aim to make room for fine-grained semantic distinctions within a structuralist framework. But there are also *skeptics* about hyperintensionality, such as Stalnaker (1984) who argues that seemingly hyperintensional distinctions can be explained away

⁷See, e.g., Stalnaker (2012, p. 758) and Stanley (2011, ch. 4).

⁸See King (2016) for an overview of structuralist approaches to semantic content.

meta-linguistically. Thus, not everyone agrees that there is a problem of hyperintensionality to begin with. While we believe that there are serious problems with both the structuralist and the skeptic stance on hyperintensionality, this is not the place to air these misgivings.⁹ Our aim here is not to defend the worlds-based approach from its structuralist critics, nor to convince the skeptic that hyperintensional phenomena should indeed find their way into our semantic theories. Rather, we aim to give those who already find the worlds-based approach attractive a way to solve the old and the new problem of hyperintensionality. To the extent that we succeed, this might be taken to confer at least some indirect support on the worlds-based approach to semantics more generally.

With these preliminary remarks in place, here is a brief overview of the paper. In §2, we begin by clarifying and refining various aspects of Frege's equipollence principle. In §3, we then use the refined equipollence principle to define a class of 'non-trivial' hyperintensions. In §4, we use these hyperintensions to give a formal rendering of the equipollence principle within a worlds-based framework. In §5, we show that the resulting semantic framework allows us to solve both the old and the new problem of hyperintensionality. In §6, we argue that the proposed framework is superior to another hyperintensional framework, due to Mark Jago (2014), which also promises to accommodate a hyperintensional, yet non-trivial notion of semantic content. In §7, we offer some concluding remarks.

2 The Fregean thesis

The aim of this section is to clarify and refine various aspects of the following generic formulation of Frege's equipollence principle:

(Fregean thesis₁) Two sentences A and B have the same content if and only if A and B are immediately recognizable as equivalent.

As mentioned, (Fregean thesis₁) offers an individuation of semantic content that is distinctively *cognitive* or *epistemic* in nature. In order for two sen-

⁹For detailed criticism, see Stanley (2010) and Jago (2014, ch. 2-3).

tences to have the same content, it is not enough for them to be equivalent—they must be so in a “trivial” or “immediate” way. And what counts as immediate or trivial is closely related to the cognitive resources that agents have available for inferential reasoning. If we focus on the Kurt Gödels of this world, even fairly complex mathematical and logical equivalences may be immediate or trivial, whereas this is not so if we focus on people as they are most. So, according to (Fregean thesis₁), the question of whether two sentences differ in content turns on the cognitive resources that agents have available. To reflect this epistemic aspect of Frege’s principle in our terminology, we will henceforth talk about the *epistemic* content of linguistic expressions.¹⁰

The epistemic aspect of Frege’s equipollence principle also reveals that (Fregean thesis₁) cannot strictly be the right way of formulating the principle. According to (Fregean thesis₁), whether two sentences have the same content depends on whether they are immediately recognizable as equivalent. However, while the relation of having the same content is most naturally understood as an identity relation, which is transitive, the relation of being immediately recognizable is *not* transitive. Here is a simple illustration of why: consider a sequence A_0, A_1, \dots, A_n of sentences such that A_n can be inferred from A_0 in no less than n inferential steps (and A_0 can likewise be inferred from A_n in no less than n steps). In other words, the shortest inference from A_0 to A_n (and *vice versa*) involves n applications of simple inference rules such as modus ponens, conjunction elimination, and the like.¹¹

¹⁰Note that Frege’s equipollence principle does not make content individuation depend on the stock of empirical information that an agent has available. Obviously, what is “immediately recognizable as equivalent” strictly speaking depends on one’s empirical information. For example, an astronomer might immediately recognize that “Hesperus is Phosphorus” is equivalent to “Hesperus is Hesperus.” But since Frege’s equipollence principle quantifies over *all* agents, regardless of their empirical background, content individuation remains insensitive to empirical information.

¹¹For example, the sequence $(A_0, A_1, A_2) = (p \wedge q, p, p \vee r)$ can be seen as a two-step derivation from $p \wedge q$ to $p \vee r$ in a proof system that includes conjunction elimination and disjunction introduction. Strictly speaking, the number of inferential steps needed to derive A_n from A_0 need not coincide with the number n , since there may be inference rules with more than one premise. For example, the sequence $(B_0, B_1, B_2) = (p \rightarrow q, p, q)$ is a one-step derivation in a proof theory that includes modus ponens. But this is immaterial for the general point that the relation of being ‘immediately recognizable as equivalent’

Given that the value of n is sufficiently high—that is, given that the shortest derivation of the equivalence of A_1 and A_n is sufficiently complex— A_1 and A_n will not count as immediately recognizable as equivalent for ordinary persons like you and me. Yet, since a single application of a simple inference rule presumably counts as immediate or obvious for agents like us, each pair of adjacent sentences A_i and A_{i+1} *will* count as immediately recognizable as equivalent.¹² So the relation of being immediately recognizable as equivalent cannot be transitive: a sentence A need not be immediately recognizable as equivalent to C , even if A is immediately recognizable as equivalent to B , and B is immediately recognizable as equivalent to C .

Thus we need to find a formulation of the equipollence principle that appeals to intransitive relations on both sides of the biconditional. This means searching for an intransitive replacement of the relation of having the *same* content that remains faithful to the epistemic aspect of the equipollence principle. For reasons that will become clear, we will refer to this relation as that of having *indistinguishable* epistemic contents. If we substitute this relation in (Fregean thesis₁), we obtain:

(Fregean thesis₂) Two sentences A and B have indistinguishable epistemic contents if and only if A and B are immediately recognizable as equivalent.

The relation of having indistinguishable epistemic contents should be understood as a term of art that captures the sense in which pairs such as (1) to (3) differ in content, whereas pairs such as (4) to (6) do not. As explained in the introduction, sentence pairs like (4) to (6) seem to say much the same thing, and seem to play much the same role in our cognitive lives. And when a pair of sentences display these features, we will say that they have ‘indistinguishable’ epistemic contents. By contrast, we will say that sentence pairs like (1) to (3), which do not display these features, have distinguishable

cannot be transitive.

¹²Of course, it is a moot point exactly what it takes for an inference rule to be “trivial” or “obvious”. For now we shall rely on an intuitive grasp of these notions, but later on, we will offer a formally precise account.

epistemic contents.¹³

A desirable consequence of individuating content by appeal to an intransitive relation is that it allows us to circumvent a recent impossibility result by Bjerring and Schwarz (2017). To see the gist of their result, consider again the sequence A_0, \dots, A_n of sentences from above. Given a high enough value of n , we can hold that A_0 and A_n differ in content for the same sorts of reasons that (3a) and (3b) differ in content. Yet, given that a single application of a simple inference rule counts as immediate or trivial, each pair of adjacent sentences A_i and A_{i+1} will *not* differ in content for the same sorts of reasons that (6a) and (6b) do not differ in content. This spells trouble for the familiar strategy of individuating content by appeal to an identity relation. For if the content of A_i is identical to that of A_{i+1} , for each i , it follows by the transitivity of identity that A_0 and A_n have identical contents as well. In other words, A_0 and A_n cannot have distinct contents, if each adjacent pair of sentences have identical contents. But as we have seen, a joint solution to the old and the new problem of hyperintensionality *should* allow us to say that A_0 and A_n differ in content (in some sense of ‘differ in content’), whereas each adjacent pair of sentences A_i and A_{i+1} do not. So, insofar as we want to develop a joint solution to the old and the new problem of hyperintensionality, Bjerring and Schwarz’s result shows that content must be individuated in an intransitive manner. And the equipollence principle promises a way to do just that.

Let us now introduce a formally precise rendering of what it means for two sentences to be ‘immediately recognizable as equivalent’. First off, note that what is ‘immediately recognizable as equivalent’ seems closely related to what is ‘immediately a priori inferable’. Consider again the sentence pairs in (4) to (6). Plausibly, these sentence pairs are immediately recognizable as equivalent precisely because it is immediately a priori inferable that they are equivalent. For example, we take it that A and $A \wedge A$ are immediately recognizable as equivalent because it is easy to establish their equivalence a

¹³To be sure, nothing of importance hinges on the exact choice of terminology. The notion of “indistinguishability” strikes us as a particularly natural way of replacing “sameness” with an intransitive counterpart, but other intuitive notions such as “cognitive synonymy” or “likeness in meaning” may do the job just as well.

priori.

What counts as immediately a priori inferable also seems to depend in the right way on the cognitive resources that agents have available. While it may be trivial for a skilled mathematician to show that $7 = 7$ is equivalent to $(5^2 \times 211 - 4) \div 753 = 7$, this may be non-trivial for people like you and me. Thus it seems natural to understand the relation of being immediately recognizable in terms of what is immediate a priori inferable.

Given that what counts as immediately a priori inferable depends on an agent's cognitive resources, we need a way of quantifying the resources that agents have available for immediate a priori reasoning. Many different measures may in principle do the job here, whether they appeal to time consumption, neural firings in the brain, or some other quantity. But to keep things relatively simple, we will think of an agent's cognitive resources in terms of the number n of steps of a priori reasoning that the agent can immediately perform. Accordingly, we will think of a priori reasoning as a *rule-bound* process whereby agents reason step-by-step by applying rules from a designated set \mathcal{S} of a priori inference rules.¹⁴

As a first stab, we can understand \mathcal{S} as an inferential system that contains enough rules to accommodate the kinds of reasoning that we find in logic, mathematics, and analytic or conceptual analysis. We are familiar with thinking about logical reasoning in terms of rules such as modus ponens, conjunction elimination, and so on. And later, in §4, we shall discuss in more detail how one might think of mathematical and conceptual reasoning in a similar rule-bound fashion. But for now, it suffices to bear two points about the system \mathcal{S} in mind. First, all rules in \mathcal{S} , whether logical or non-logical, should be strictly *a priori*. For example, there should be no rules in \mathcal{S} that allow agents to move from sentences containing 'Hesperus' to sentences containing 'Phosphorus' since it requires significant empirical information to connect facts about Hesperus to facts about Phosphorus. Second, our framework below does not presuppose any particular specification of the rules

¹⁴For similar ways of modelling resource-bounded reasoning, see Bjerring and Skipper (forthcoming), Drapkin and Perlis (1986), Jago (2014), Rasmussen (2015), and Smets and Solaki (2018).

in \mathcal{S} . For all we say, there may be a uniquely correct way of encoding a priori reasoning as a rule-bound process. But just as different contexts may require different levels of cognitive resources (and hence different values of n), different contexts will likely also require different reasoning mechanisms (and hence different specifications of \mathcal{S}). So we take it to be a desirable feature of our framework that it allows for different specifications of the system \mathcal{S} .

If a sentence B can be inferred from another sentence A within n applications of the rules in \mathcal{S} , we will say that B is “ n -step inferable in \mathcal{S} ” from A . Accordingly, we will say that B is immediately a priori inferable from A if and only if B is n -step inferable in \mathcal{S} from A . If B can be inferred from A by *some* (sufficiently large) finite number of applications of the rules in \mathcal{S} , we will say that B is *in principle* inferable in \mathcal{S} from A . And if A is n -step inferable in \mathcal{S} from the empty set, we will simply say that A is n -step inferable in \mathcal{S} .

We will make two further assumptions about n -step inferability in \mathcal{S} . First, we will assume that nothing is 0-step inferable in \mathcal{S} , except the trivial inference ‘ A therefore A ’. Nothing of importance hinges on this assumption, but it will simplify the exposition. Second, for our central result below, we will assume that \mathcal{S} is monotonic. That is, we will adopt the orthodox view that valid a priori inferences cannot be defeated by the addition of further premises. This assumption is not entirely uncontroversial, but we find it plausible enough to serve as a working assumption in this context.¹⁵

In order to reflect the step-based picture of immediate a priori inferability in our terminology, we will henceforth talk about the “epistemic n -content” of a sentence. This gives us the following rendering of the equipollence principle:

(Fregean thesis₃) Two sentences A and B have indistinguishable epistemic n -contents if and only if it is n -step inferable in \mathcal{S} that A and B are equivalent.

Let us highlight two important features of (Fregean thesis₃). First, the principle preserves the intransitivity of the previous renderings of the equipollence principle: C need not be n -step inferable from A , even if B is n -step infer-

¹⁵See Russell (2017) for additional background on defeasibility of a priori reasoning.

able from A , and C is n -step inferable from B . For example, if \mathcal{S} contains conjunction elimination, A will be 1-step inferable from $A \wedge A$, and $A \wedge A$ will be 1-step inferable from $(A \wedge A) \wedge (A \wedge A)$, but A will not be 1-step inferable from $(A \wedge A) \wedge (A \wedge A)$. So the notion of n -step inferability in \mathcal{S} provides us with an intransitive relation that remains faithful to the epistemic aspect of the equipollence principle.

Second, by varying the value of n , (Fregean thesis₃) gives rise to a whole *spectrum* of notions of content with different levels of granularity. In the limit where $n = 0$, content is individuated in a highly fine-grained manner: any two sentences will come out with distinguishable epistemic n -contents. In the opposite limit where n goes towards infinity, content is individuated in a highly coarse-grained manner: any two a priori equivalent sentences will come out with indistinguishable epistemic n -contents. For intermediate values of n , we find notions of content that are hyperintensional, yet non-trivial: some, but not all, a priori equivalent sentences will have indistinguishable epistemic n -contents.

With a refined version of the equipollence principle at our disposal, the next task is to develop a worlds-based semantic framework that satisfies (Fregean thesis₃).

3 Epistemic n -intensions

The guiding idea behind our semantic framework is to understand the notion of ‘ n -step inferability in \mathcal{S} ’ as a kind of epistemic necessity, which is sensitive to an agent’s cognitive resources. More specifically, we will say that a sentence A is necessary (in the relevant epistemic sense of ‘necessity’) if and only if A is n -step inferable in \mathcal{S} . Hence, since only a priori truths are in principle inferable in \mathcal{S} , only a priori truths will ever be necessary in our framework. But *which* truths come out as necessary depends on the value of n . When $n = 0$, *no* a priori truths come out as necessary, since we assume that nothing is 0-step inferable in \mathcal{S} . When n approaches infinity, *all* a priori truths come out as necessary, since we assume that all a priori truths are in principle inferable in \mathcal{S} . For intermediate values of n , some, but not all, a

priori truths come out as necessary.

Why understand n -step inferability in \mathcal{S} as a kind of necessity? Because it allows us to give a semantic analysis of (Fregean thesis₃) that mirrors the following classical individuation principle from intensional semantics:

(Intensional content) Two sentences A and B have the same intensional content if and only if the intension of $A \leftrightarrow B$ is necessary.

In our framework, the relevant epistemic notion of necessity will play much the same role as the standard non-epistemic notion of necessity plays in (Intensional content). More specifically, we will define a class of non-trivial hyperintensions—or, as we shall call them, ‘epistemic n -intensions’—that will allow us to say that A and B have indistinguishable epistemic n -contents if and only if the epistemic n -intension of $A \leftrightarrow B$ is necessary.

To situate our proposal within the existing theoretical landscape, it is instructive briefly to compare it to the general approach that we find in Chalmers’ epistemic two-dimensional semantics.¹⁶ Roughly, on this account, two sentences are said to have the same *epistemic* content if and only if they are a priori equivalent. As in our framework, Chalmers’ framework is based on an understanding of apriority in terms of a notion of epistemic necessity. But in contrast to the notion of *immediate* a priori inferability that we are targeting, Chalmers targets an *idealized* notion of apriority, which is not sensitive to an agent’s cognitive resources. As a result, his framework renders all a priori truths epistemically necessary, and hence associates all a priori truths with the same content. However, Chalmers also speculates that, if we want to cut content more finely than a priori equivalence, we might try to develop a semantics around a *non-ideal* notion of epistemic necessity, which *is* sensitive to limited cognitive resources. Among his candidates for such a non-ideal notion of epistemic necessity are “knowability through such-and-such amount of a priori reasoning” and “provability in n steps of logical reasoning” (Chalmers 2011, pp. 102-03). Our proposal below may be seen as an attempt to put flesh on these bones.

As a first step, we need to define a space of worlds or ‘scenarios’ at which

¹⁶See, e.g., Chalmers (2006; 2011).

we can evaluate sentences for truth and falsity. We will think of scenarios as representations of ways the world might be for all an agent can tell a priori. Given that some agents have limited resources available for a priori reasoning, some scenarios will give an a priori inconsistent representation of the world. Such scenarios look possible to agents with limited capacities for a priori reasoning, but are nonetheless impossible in the sense that it is *in principle* possible to reveal them as being a priori inconsistent. Moreover, given that some (idealized) agents have unlimited resources available for a priori reasoning, some scenarios will give an a priori consistent representation of the world. Such scenarios are a priori possible in the sense that no amount of a priori reasoning will reveal them as being inconsistent.

For present purposes, we will construct scenarios linguistically as sets of sentence types in some world-making language \mathcal{L} . While there is much to say about the specifics of such a world-making language, it suffices here to specify just a few general requirements that \mathcal{L} should meet.¹⁷ First, the content of sentence types in \mathcal{L} should be sufficiently fine-grained and, in particular, more fine-grained than Fregean content. For if the content of sentences in \mathcal{L} is more fine-grained than Fregean content, there should be no obvious circularities involved in appealing to scenarios constructed in \mathcal{L} to develop a semantics built around Frege’s equipollence principle. Second, \mathcal{L} should be rich enough to represent every possible and impossible way the world might be for agents with (potentially) limited resources available for a priori reasoning. In particular, it should allow us to express the kinds of sentences that we find in logic, mathematics, science, and so on. Finally, \mathcal{L} should display no context-dependent features. For if we constructed scenarios in a context-dependent language, scenarios need not give determinate representations of ways the world might (or might not) be. For instance, if a scenario included a demonstrative sentence like “this is blue,” it would be unclear what this scenario represented, and hence unclear what would be true and false at that scenario. But aside from such general requirements, we need not make any substantive assumptions about the world-making language \mathcal{L} .

¹⁷For further considerations in this direction, see Berto (2013), Chalmers (2011), and Jago (2014).

Next, we need to say what it means for a sentence token in English to be true or false at a scenario. For ease of exposition, we shall ignore the relation between English and the world-making language \mathcal{L} and simply pretend that scenarios correspond to sets of sentences in English. This allows us to evaluate sentence tokens in English—henceforth just ‘sentences’—for truth and falsity at scenarios by reference to simple set-membership.¹⁸ That is, if V is an evaluation function from sentences and scenarios to truth-values:

(Verification) $V(A, w) = T$ if and only if $A \in w$.

(Falsification) $V(A, w) = F$ if and only if $\neg A \in w$.

If w neither verifies nor falsifies A , we will say that A is *indeterminate* at w and that w contains a ‘truth-value gap’.

While every scenario corresponds to a set of sentences, not all sets of sentences should correspond to a scenario. For recall that we want scenarios to represent ways the world might be for agents with (potentially) limited resources available for a priori reasoning. And there are non-trivial constraints on what the world might be like for such agents. In particular, scenarios that are “blatantly” impossible in the sense that it is trivial or immediate for a given agent to see that they are inconsistent do not count as ways the world might be for the agent.¹⁹ Thus, given that we focus on agents who can immediately or trivially perform n steps of a priori reasoning, we will restrict our attention to scenarios from which no hidden contradiction can be inferred within n steps of a priori reasoning. That is, we will restrict our attention to scenarios that are *n-consistent* in the following sense (where ‘ $\Gamma \vdash_{\mathcal{S}}^n \Gamma'$ ’ means that the set Γ' of sentences can be inferred from the set Γ of sentences by n applications of the rules in \mathcal{S}):

(n-consistency) A scenario w is *n-consistent* if and only if $w \not\vdash_{\mathcal{S}}^n \perp$, where \perp is an arbitrary contradiction of the form $\{A, \neg A\}$.

¹⁸A detailed account of the relation between \mathcal{L} and English would appeal to a translation of tokens in English to types in \mathcal{L} ; see Jago (2014, §5) for such an account.

¹⁹This sort of requirement goes back at least to Hintikka (1975) who noted that “impossible possible worlds” may well be epistemically possible for ordinary agents as long as they are not too logically ill-behaved; see also Lewis (2004), Bjerring (2013), and Jago (2014) for similar ideas.

In other words, a scenario w is n -inconsistent if and only if the shortest derivation of a contradiction from w is no longer than n steps in \mathcal{S} .²⁰ Here is a simple illustration: suppose that \mathcal{S} contains just modus ponens, and let $w = \{A, A \rightarrow B, \neg B\}$. Since it takes exactly one application of modus ponens to derive a contradiction from w , w is 0-consistent, but 1-inconsistent. By comparison, the scenario $v = \{A, A \rightarrow B, B \rightarrow C, \neg C\}$ is 1-consistent, but 2-inconsistent, since it takes exactly two applications of modus ponens to derive a contradiction from v . Finally, a scenario such as $u = \{A \vee B\}$ is consistent *simpliciter* since it is not possible to derive a contradiction from u in any number of steps in \mathcal{S} .

Henceforth, let W_n denote the set of all n -consistent scenarios. Since all scenarios are at least 0-consistent, no scenario in W_n will ever verify an explicit contradiction of the form $\{A, \neg A\}$. In other words, while scenarios in W_n might well contain truth-value *gaps*, they will never contain truth-value *gluts*. By varying the value of n , we can vary the ‘degree of consistency’ of the scenarios in W_n . For $n = 0$, W_n is a highly permissive space of scenarios, which excludes only scenarios that verify explicit contradictions. By contrast, when n approaches infinity, W_n becomes highly restrictive, excluding all scenarios that fail to be a priori consistent *simpliciter*. For intermediate values of n , the scenarios in W_n will display different intermediate degrees of a priori consistency.

While (n -consistency) guarantees that scenarios in W_n are not too blatantly inconsistent, it does not itself give us the required tools to treat n -step inferability as a kind of necessity. For on the usual understanding necessity, a sentence is necessary if and only if it is true across all worlds (in some designated space of worlds). But since (n -consistency) does not impose any constraints on how many truth-value gaps scenarios in W_n may contain, *no* sentence will ever be true at all scenarios in W_n , regardless of the value of n . To ensure that our framework accommodates an appropriate class of necessities, we therefore need a formal device that can help us to eliminate certain

²⁰Accordingly, whether a contradiction can be derived from w in *more* than n steps is irrelevant for whether w is n -consistent. Thanks to an anonymous reviewer for asking us to clarify this point.

obvious or blatant truth-value gaps from scenarios in W_n :

(n -expansion) For any scenario $w \in W_n$, the n -expansion of w is written ' w^n ' and defined as: $w^n = \{A \mid w \vdash_{\mathcal{S}}^n A\}$.

In other words, the n -expansion of w contains all sentences that can be inferred from w within n steps in \mathcal{S} (including those sentences that are already contained in w). For example, suppose that \mathcal{S} contains just conjunction elimination and double negation elimination, and let $w = \{\neg\neg(A \wedge B)\}$. The 1-expansion of w is then given by $w^1 = \{\neg\neg(A \wedge B), A \wedge B\}$, whereas the 2-expansion of w is given by $w^2 = \{\neg\neg(A \wedge B), A \wedge B, A, B\}$. Note that the 0-expansion of w is identical to w itself, since nothing is 0-step inferable in \mathcal{S} . Note also that the n -expansion of scenarios in W_n will never verify explicit contradictions, since all scenarios in W_n are n -consistent.

We are now in a position to define the *epistemic n -intension* of a sentence as a function from scenarios in W_n to truth-values:

(Epistemic n -intension) The epistemic n -intension of A is true at w if and only if $V(A, w^n) = T$, and false at w if and only if $V(A, w^n) = F$.

As usual, if the epistemic n -intension of A is neither true nor false at w , we will say that A is *indeterminate* at w .

By varying the value of n , we get a whole spectrum of epistemic n -intensions. When $n = 0$, epistemic n -intensions will be highly a priori ill-behaved: they are only guaranteed not to verify and falsify the same sentence at the same scenario. By contrast, when n approaches infinity, epistemic n -intensions will be highly a priori well-behaved: they will deliver truth-values that are in full accordance with what can in principle be established a priori about our world. For intermediate values of n , epistemic n -intensions will display different degrees of a priori well-behavedness.

Finally, we can define what it means for epistemic n -intensions to be possible or necessary:

(Possibility) The epistemic n -intension of A is possible if and only if the epistemic n -intension of A is true at some scenario in W_n .

(Necessity) The epistemic n -intension of A is necessary if and only if the epistemic n -intension of A is true at all scenarios in W_n .

If the epistemic n -intension of A is possible but not necessary, we will say that A is *contingent*. As in other frameworks that allow for truth-value gaps, possibility and necessity cannot be treated as duals, since the epistemic n -intension of A need not be true at w , even if it fails to be false at w (and *vice versa*). This is why possibility and necessity are defined separately.

Despite obvious differences, epistemic n -intensions are comparable to classical intensions in a number of respects. Here are three notable examples. First, classical intensions are often understood as “carving up” the space of possible worlds into those worlds that verify a given sentence, and those that do not. Similarly, we can understand epistemic n -intensions as carving up the space of n -consistent scenarios into those scenarios whose n -expansions verify a given sentence, and those whose n -expansions do not. Second, classical intensions allow us to spell out a notion of necessity in terms of intensions being *true-at-all-possible-worlds*. Similarly, we have seen that epistemic n -intensions can be used to spell out an epistemic notion of necessity in terms of epistemic n -intensions being *true-at-all-worlds-in- W_n* . Finally, classical intensions are often used to give set-theoretic representations of various logical connectives. For example, the meaning of conjunction (‘ \wedge ’) and disjunction (‘ \vee ’) can be expressed by the following subset relations (where ‘ $|A|$ ’ denotes the set of possible worlds that verify A):

$$\begin{array}{ll} (\wedge) & |A \wedge B| \subseteq |A| \cap |B| \\ & |A| \cap |B| \subseteq |A \wedge B| \\ (\vee) & |A \vee B| \subseteq |A| \cup |B| \\ & |A| \cup |B| \subseteq |A \vee B| \end{array}$$

It is well-known that these classical dependencies fail once we move into an impossible-worlds framework.²¹ However, we can still use epistemic n -intensions to give informative set-theoretic representations of various logical connectives. For example, if \mathcal{S} contains the standard introduction and elimination rules for conjunction and disjunction, we get the following subset

²¹For details, see for instance Elliott (2017).

relations (where $|A|^n$ denotes the set of scenarios in W_n at which the epistemic n -intension of A is true):²²

$$\begin{array}{ll}
 (\wedge) & |A \wedge B|^n \subseteq |A|^{n+1} \cap |B|^{n+1} \\
 & |A|^n \cap |B|^n \subseteq |A \wedge B|^{n+1} \\
 (\vee) & |A \vee B|^n \cap |\neg A|^n \subseteq |B|^{n+1} \\
 & |A|^n \subseteq |A \vee B|^{n+1}
 \end{array}$$

These subset relations show that, although not truth-functional in the classical sense, our semantics nevertheless offers informative dependencies between the epistemic n -intensions of complex sentences and those of their parts.

With the core elements of our semantic framework on the table, we can now return to Frege's equipollence principle.

4 The Fregean thesis formalized

Just as the intensional theory identifies the semantic content of a sentence with its intension, we will identify the epistemic n -content of a sentence with its epistemic n -intension. Accordingly, we will accommodate Frege's equipollence principle within our semantic framework in the following way:

(Fregean thesis) Two sentences A and B have indistinguishable epistemic n -contents if and only if the epistemic n -intension of $A \leftrightarrow B$ is necessary.

As promised, this principle offers an intransitive individuation of epistemic n -content: the epistemic n -intension of $A \leftrightarrow C$ need not be necessary, even if the epistemic n -intensions of $A \leftrightarrow B$ and $B \leftrightarrow C$ are both necessary. For example, if \mathcal{S} contains the standard introduction and elimination rules for conjunction, the epistemic 2-intensions of $A \leftrightarrow (A \wedge A)$ and $(A \wedge A) \leftrightarrow (A \wedge A) \wedge (A \wedge A)$ will both be necessary, whereas the epistemic 2-intension of $A \leftrightarrow (A \wedge A) \wedge (A \wedge A)$ will not.

²²For those interested in the formal details, here is a brief proof of the upper subset relation involving conjunction: suppose $w \in |A \wedge B|^n$. By (Epistemic n -intension) and (Verification), $A \wedge B \in w^n$. Given that \mathcal{S} contains conjunction elimination, $w^n \vdash_{\mathcal{S}}^1 A$ and $w^n \vdash_{\mathcal{S}}^1 B$. Hence, $w \vdash_{\mathcal{S}}^{n+1} A$ and $w \vdash_{\mathcal{S}}^{n+1} B$. By (n -expansion), $A \in w^{n+1}$ and $B \in w^{n+1}$. By (Epistemic n -intension) and (Verification), the epistemic $n+1$ -intension of A is true at w , and similarly for B . Thus, $w \in |A|^{n+1}$ and $w \in |B|^{n+1}$, which means that $w \in |A|^{n+1} \cap |B|^{n+1}$. The other subset relations can be established in similar ways.

The following result shows that (Fregean thesis) satisfies (Fregean thesis₃) in the sense that the two theses deliver precisely the same semantic verdicts:

Theorem 1. The epistemic n -intension of A is necessary if and only if $\vdash_{\mathcal{S}}^n A$.

Proof. For left-to-right, suppose that the epistemic n -intension of A is necessary. By (Necessity), the n -intension of A is true at every scenario $w \in W_n$. By (Epistemic n -intension), $V(A, w^n) = T$, for every $w \in W_n$. So, by (Verification), $A \in w^n$, for all $w \in W_n$. Since $\emptyset \in W_n$, $A \in \emptyset^n$. So, by (n -expansion), $\vdash_{\mathcal{S}}^n A$. For right-to-left, suppose that $\vdash_{\mathcal{S}}^n A$. By monotonicity of \mathcal{S} , $w \vdash_{\mathcal{S}}^n A$, for all $w \in W_n$. So, by (n -expansion), $A \in w^n$, for all $w \in W_n$. By (Verification), $V(A, w^n) = T$, for every $w \in W_n$. By (Epistemic n -intension), the epistemic n -intension of A is true at every scenario $w \in W_n$. Thus, by (Necessity), the epistemic n -intension of A is necessary. \square

It follows as an immediate corollary of Theorem 1 that the epistemic n -intension of $A \leftrightarrow B$ is necessary if and only if $A \leftrightarrow B$ is n -step inferable in \mathcal{S} . Hence, (Fregean thesis) and (Fregean thesis₃) individuate epistemic n -content in precisely the same manner.

By varying the value of n , we can use (Fregean thesis) to reason about a whole spectrum of notions of content with different levels of granularity. When $n = 0$, content is individuated in a highly fine-grained manner: any two sentences are associated with distinguishable epistemic n -contents. By contrast, when n goes towards infinity, content is individuated in a highly coarse-grained manner: any two a priori equivalent sentences are associated with indistinguishable epistemic n -contents. For intermediate values of n , we get a range of notions of hyperintensional, yet non-trivial content: some, but not all, a priori equivalent sentences are associated with indistinguishable epistemic n -contents.

There are three interesting observations to make about the limit where n approaches infinity. First, while the resulting notion of content is highly coarse-grained, it remains hyperintensional in the sense that *a posteriori* equivalent sentences are still associated with distinguishable epistemic n -contents. For example, ‘Hesperus is Hesperus’ and ‘Hesperus is Phosphorus’

have distinguishable epistemic n -contents, for any value of n , since \mathcal{S} does not contain any rules that allow agents to move from facts about Hesperus to facts about Phosphorus (or *vice versa*).

Second, the relation of having indistinguishable epistemic n -content turns into an identity relation in the limit where n approaches infinity. For in that case, what is n -step inferable in \mathcal{S} coincides with what is *in principle* inferable in \mathcal{S} . Hence, if A and B are a priori equivalent, the epistemic n -intension of $A \leftrightarrow B$ is necessary, when n goes towards infinity. As a result, the epistemic n -intensions of A and B will be co-extensional at all scenarios in W_n .

Third, when n goes towards infinity, epistemic n -intensions individuate content in much the same way as Chalmers' so-called *primary intensions*.²³ On Chalmers' picture, the primary intension of a sentence represents what we might call its "ideal" epistemic content: roughly, the conditions under which the sentence is true or false at a priori closed and consistent scenarios. Two sentences A and B are then said to have the same ideal epistemic content if and only if the primary intension of $A \leftrightarrow B$ is necessary, which holds if and only if it is a priori true for an agent with no cognitive limitations that A and B are equivalent (Chalmers 2011). Essentially, that is, A and B have the same ideal epistemic content if and only if it is in principle inferable in \mathcal{S} that they are equivalent. As we saw above, when $A \leftrightarrow B$ is in principle inferable in \mathcal{S} , A and B will have indistinguishable epistemic n -contents, for n going towards infinity. So epistemic n -intensions behave much like primary intensions in the limit where n approaches infinity. For finite values of n , we may think of epistemic n -intensions as "non-ideal" primary intensions that represent the "non-ideal" epistemic content of sentences: roughly, the conditions under which sentences are true or false at the n -expansions of scenarios in W_n .²⁴ As such, there is a natural sense in which epistemic n -intensions may be understood as generalized Chalmersian primary intensions.

²³See Chalmers (2002; 2011) for details.

²⁴Of course, these similarities between epistemic n -intensions and primary intensions presuppose that \mathcal{S} contains enough rules to allow us to infer the class of sentences that count as *a priori* in Chalmers' framework.

5 Revisiting the two problems of hyperintensionality

We now want to show how the proposed semantic framework can provide a joint solution to the old and the new problem of hyperintensionality. To do so, we must look for a value of n for which the sentence pairs in (1) to (3) get associated with distinguishable epistemic n -contents, whereas the sentence pairs in (4) to (6) get associated with indistinguishable epistemic n -contents. Obviously, the limit cases where $n = 0$ or n approaches infinity are non-starters: when n approaches infinity, we cannot solve the *old* problem of hyperintensionality because all a priori equivalent sentences get associated with indistinguishable epistemic n -contents; and when $n = 0$, we cannot solve the *new* problem of hyperintensionality because all sentences whatsoever get associated with distinguishable epistemic n -contents. What is needed is an *intermediate* value of n . The appropriate value is, of course, going to depend on the exact specification of the rules in \mathcal{S} . But for the sake of concreteness, let us illustrate how our semantic framework can deliver the intuitively correct verdicts in (1) to (6) when $n = 2$ and \mathcal{S} contains just a few familiar inference rules.

Consider first the logical cases (3) and (6):

- (3) a. A .
 b. $(A \wedge (\neg(\neg A \vee B))) \vee (A \wedge B)$.
- (6) a. A .
 b. $A \wedge A$.

Given that \mathcal{S} contains the standard introduction and elimination rules for conjunction and disjunction, $A \wedge A$ is 1-step inferable in \mathcal{S} from A , and A is 1-step inferable in \mathcal{S} from $A \wedge A$. So $A \leftrightarrow (A \wedge A)$ is 2-step inferable in \mathcal{S} . By Theorem 1, this means that the epistemic 2-intension of $A \leftrightarrow (A \wedge A)$ is necessary.²⁵ Hence, by (Fregean thesis), it follows that A and $A \wedge A$ have

²⁵Again, depending on how we exactly specify the logical fragment of \mathcal{S} , further steps might be needed to infer $A \leftrightarrow (A \wedge A)$. For example, if we adopt a Lemmon-style natural deduction system, it would take five proof steps to infer $A \leftrightarrow (A \wedge A)$, because we would need additional steps for making and discharging assumptions. But such details are unimportant for the issue at hand.

indistinguishable epistemic 2-contents. By contrast, since there are no logical rules that allow us to infer the complex biconditional $A \leftrightarrow (A \wedge (\neg(\neg A \vee B))) \vee (A \wedge B)$ in just 2 steps, it will not be 2-step inferable in \mathcal{S} that (3a) and (3b) are equivalent. Thus, by Theorem 1, (3a) and (3b) have distinguishable epistemic 2-contents. So (Fregean thesis) delivers the right semantic verdicts in the logical cases.

Consider next the mathematical cases (2) and (5):

- (2) a. $7 = 7$.
- b. $(5^2 \times 211 - 4) \div 753 = 7$.
- (5) a. $x = y$.
- b. $y = x$.

Given that \mathcal{S} contains a rule for commutativity of identity that allows us to infer $x = y$ from $y = x$ in one step, it follows that $x = y \leftrightarrow y = x$ is 2-step inferable in \mathcal{S} . By Theorem 1 and (Fregean thesis), this means that $x = y$ and $y = x$ have indistinguishable epistemic 2-contents. By contrast, $7 = 7$ and $(5^2 \times 211 - 4) \div 753 = 7$ have distinguishable epistemic 2-contents, since no reasonable specification of the mathematical fragment of \mathcal{S} will make it possible to infer $7 = 7 \leftrightarrow (5^2 \times 211 - 4) \div 753 = 7$ in just two steps. So (Fregean thesis) delivers the right semantic verdicts in the mathematical cases.

Finally, consider the analytic and metaphysical cases (1) and (4):

- (1) a. Hesperus is Hesperus.
- b. Hesperus is Phosphorus.
- (4) a. The number of primes is infinite.
- b. There are infinitely many primes.

As previously mentioned, ‘Hesperus is Hesperus’ and ‘Hesperus is Phosphorus’ get associated with distinguishable epistemic n -contents, for *any* value of n , since \mathcal{S} only contains a priori inference rules. So (Fregean thesis) delivers the right semantic verdict in the metaphysical case. As for the analytic case, we can ensure that ‘the number of primes is infinite’ and ‘there are infinitely many primes’ get associated with indistinguishable epistemic 2-contents by

including in \mathcal{S} a simple a priori rule that allows us to infer that ‘the number of F s is x ’ from ‘there are x many F s’ (and *vice versa*). So (Fregean thesis) delivers the right semantic verdict in the analytic case as well.

This illustrates how our semantic framework can provide a joint solution to the old and the new problem of hyperintensionality. It solves the old problem of hyperintensionality by allowing us to draw semantic distinctions between sentence pairs—such as (1) to (3)—that are equivalent, but only non-trivially so. And it solves the new problem of hyperintensionality by allowing us to ban semantic distinctions between sentence pairs—such as (4) to (6)—that are trivially equivalent. In other words, the proposed semantic framework offers a way of making cuts in content that are neither implausibly coarse-grained nor implausibly fine-grained.

While we have illustrated on a case-by-case basis how we can specify the rules in \mathcal{S} so as to ensure that (Fregean thesis) delivers the right semantic verdicts, it remains an open question whether it is feasible to model all sound a priori reasoning in a step-wise fashion. Obviously, there is no requirement that the rules in \mathcal{S} need be anywhere near as simple as the ones we made use of above. Nor is there any requirement that \mathcal{S} contains only one type of rule for sentences containing a given expression. The rules in \mathcal{S} need not even offer a *positive* characterization of the types of situations in which a given sentence can be inferred from a set of premises. It may turn out that, for some sentences, we can only specify rules that allow us to infer their negations. For instance, while Gettier cases have taught us a lot about situations in which it is *not* correct to say that a subject knows some proposition, it remains unclear whether there is a specifiable set of conditions under which it is always correct to say that the subject *does* know some proposition.

More generally, the issues involved in specifying \mathcal{S} are very similar to issues that arise in discussions of entailment patterns in natural languages. If we adopt a too restrictive view on what entailment is—say, entailment in propositional logic—we will fail to capture many entailment patterns in natural language. For instance, while a sentence like “Peter painted the fence” lexically entails that “a fence was painted”, the former does not *logically* entail the latter. So to capture such entailments, we must focus on a broader,

more permissive notion of entailment. Questions concerning how to specify this broader notion of entailment bear resemblance to questions concerning how we should specify the rules in \mathcal{S} to capture a sufficiently permissive notion of a priori entailment. As such, it might be useful to look to work in formal semantics when deciding which rules should be included in \mathcal{S} . For instance, we might want \mathcal{S} to contain an appropriate set of *meaning postulates* similar in spirit to those that semanticists in the Montagovian tradition have used to define a notion of entailment that is less strict than logical entailment. Or we might want \mathcal{S} to contain various rule-bound encodings of semantic relations that exist among entries in a natural language lexicon.²⁶

But regardless of whether it is feasible to capture all sound a priori reasoning in this way, let us stress that the guiding idea behind our semantics does not rest on this assumption. As explained in §3, we are free to use any quantitative measure of the complexity of different lines of a priori reasoning; and we take it to be uncontroversial that different lines of a priori reasoning indeed *do* differ in complexity.

6 Jago on Fregean content

In recent work, Jago (2013b; 2014) has developed a hyperintensional framework that—much like ours—promises to accommodate a notion of hyperintensional, yet non-trivial content. In this section, we briefly review his framework and argue that ours is superior in a number of respects.

Jago’s hyperintensions are defined over a set of possible and impossible worlds—call it $W_{\mathcal{J}}$ —where the impossible worlds are assumed not to contain any “blatant” or “trivial” inconsistencies. According to Jago, it is a *vague* matter what counts as a trivial inconsistency, and hence a vague matter which impossible worlds belong to $W_{\mathcal{J}}$. For example, while a world that falsifies the trivial logical truth “ $1 + 3 = 3 + 1$ ” is determinately not part of $W_{\mathcal{J}}$, it is indeterminate whether worlds that falsify more subtle logical truths are part of $W_{\mathcal{J}}$ (Jago 2014, ch. 7). Yet, irrespective of how the line is drawn between trivial and non-trivial inconsistencies, Jago allows that worlds in $W_{\mathcal{J}}$ may be

²⁶For further discussion, see Glanzberg (2015) and Zimmermann (1999).

arbitrarily partial or “truth-gappy”. That is, while no world in $W_{\mathcal{J}}$ *falsifies* trivial logical truths such as $A \vee \neg A$, there are plenty of worlds in $W_{\mathcal{J}}$ that fail to *verify* such truths.

Since the boundaries of $W_{\mathcal{J}}$ are vague, Jago’s account of semantic content treats content as a vague phenomenon. More specifically, on Jago’s account, two sentences A and B do not *determinately* differ in content if and only if no world that is determinately part of $W_{\mathcal{J}}$ verifies A and falsifies B , and *vice versa* (Jago 2014, p. 255). For instance, since no world that verifies $A \wedge B$ and falsifies $B \wedge A$ is determinately part of $W_{\mathcal{J}}$, $A \wedge B$ and $B \wedge A$ do not determinately differ in content. By contrast, if B is a complicated logical equivalent of A , A and B *do* determinately differ in content. For there are worlds that determinately belong to $W_{\mathcal{J}}$, which verify A and falsify B . Subtleties aside, this is how Jago proposes to treat semantic content as a vague phenomenon.

But why should we understand semantic content as a vague phenomenon in the first place? Jago’s motivation stems from an alleged vagueness at the level of *knowledge ascriptions*, which he tries to capture by developing an epistemic logic that satisfies the following principle (the label is ours):

(Epistemic indeterminacy) If ‘ A ’ is a trivial consequence of what an agent i knows, then it’s never determinate that i fails to know that A . (Jago 2013b, pp. 1166-1167)

According to Jago, this principle helps us solve a resilient problem within epistemic logic: the problem of modeling boundedly rational epistemic agents; that is, agents who fall short of logical perfection or “logical omniscience”, but who nevertheless display some amount of logical competence. Consider an agent who is *minimally* logically competent in the sense that the agent knows at least all the *trivial* logical consequences of what she knows. The problem is that it can be shown that such agents must thereby also know *all* logical consequences of what they know, and hence count as logically omniscient. Jago refers to this problem as “the problem of rational knowledge” (Jago 2013b).²⁷ To avoid logical omniscience, Jago concludes that it must

²⁷For further discussion of this result, see Bjerring (2013) and Bjerring and Skipper

be “possible to fail to know or believe trivial truths (and, more generally, trivial consequences of one’s beliefs” (Jago 2014, p. 243). Yet, says Jago, if agents can miss out on trivial logical consequences of what they know, it seems that we have merely traded logical omniscience for radical logical incompetence. So to retain an appropriate measure of logical competence, Jago submits that it cannot be *determinate* that a given agent fails to know any particular trivial consequence of what she knows. This is the motivation behind (Epistemic indeterminacy).

Given (Epistemic indeterminacy), it follows that whenever A and B are trivially equivalent, any agent who knows A cannot determinately fail to know B , and *vice versa*. If we combine this result with the claim that any two trivially equivalent sentences A and B do not determinately differ in content, we thus obtain the following thesis: two sentences A and B do not determinately differ in content if and only if anyone who knows A does not determinately fail to know B , and *vice versa*. While Jago does not explicitly state this relationship between his models of knowledge and semantic content, it seems to capture his motivation for treating semantic content as a vague phenomenon.

However, Jago’s reliance on (Epistemic indeterminacy) in the context of epistemic logic faces a number of problems, and since his motivation for treating semantic content as a vague phenomenon derives from (Epistemic indeterminacy), these problems carry over to the semantic context as well. Elsewhere, we have leveled a detailed critique of (Epistemic indeterminacy) and its role in Jago’s epistemic logic.²⁸ But it is worth here highlighting three problems that we find particularly troublesome for his semantic theory.

First, we remain unconvinced by Jago’s central motivation for (Epistemic indeterminacy), which trades on a purported analogy between the problem of rational knowledge and the Sorites Paradox. Here is Jago’s motivation of the analogy:

The problem of rational knowledge can be formulated in terms of a step-by-step deduction D from premises C which the agent in question

(forthcoming).

²⁸Bjerring and Skipper (forthcoming, §4).

clearly knows, to a conclusion ‘ A ’ that the agent clearly does not know. By assumption, not every step of reasoning in D preserves the agent’s knowledge (since we eventually arrive at a conclusion the agent does not know). Yet any attempt to say precisely at which point in the deduction the agent’s knowledge gives out is doomed to failure. [...] Formulating the problem in this way brings out its structural similarity with the sorites paradox. In this case, the principle that rational agents know the trivial consequences of what they know plays the role that tolerance conditionals play in the sorites. (Jago 2013b, p. 1155)

Our worry is that the alleged analogy between the problem of rational knowledge and the Sorites Paradox in fact presupposes (Epistemic indeterminacy), in which case it cannot provide non-circular support for (Epistemic indeterminacy). Recall that Jago motivates the analogy by saying that “any attempt to say precisely at which point in the deduction the agent’s knowledge gives out is doomed to failure.” Yet, this claim seems plausible only if we already accept (Epistemic indeterminacy). Otherwise, nothing would prevent us from saying precisely at which point in the relevant deduction the agent’s knowledge gives out. As such, the analogy seems to presuppose rather than support (Epistemic indeterminacy). Consequently, Jago has offered no independent motivation for (Epistemic indeterminacy), and hence no independent motivation for treating semantic content individuation as a vague phenomenon.

Second, even if we grant that semantic content may be treated as a vague phenomenon, Jago’s account does not seem to deliver the semantic verdicts that we should expect of such an account. On his account, we know that whenever A and B are trivially equivalent, A and B do not determinately differ in content. But in many cases, it seems, it is simply too weak only to say that A and B *do not determinately* differ in content. For example, it seems obviously determinate that pairs such as ‘ $x = y$ ’ and ‘ $y = x$ ’ do not differ in content, just as it seems obviously determinate that ‘ $7 = 7$ ’ and ‘ $(5^2 \times 211 - 4) \div 753 = 7$ ’ do. More generally, as we hope to have motivated throughout the paper, whenever two sentences A and B are trivially equivalent, we are prepared to assert outrightly that they do not differ in content. But then it

also seems clear that we are prepared to assert that it is determinate that A and B do not differ in content. Yet, Jago’s framework does not deliver this result. It only delivers the weak result that trivially equivalent sentences do not determinately differ in content.

Finally, since worlds in $W_{\mathcal{J}}$ may contain arbitrarily many truth-value gaps, no sentence will ever be necessary in the sense of being true across all worlds in $W_{\mathcal{J}}$. This strikes us as an odd, if not straightforwardly problematic result. Since Carnap (1947) we have been used to thinking about content individuation in terms of different notions of necessity. On Carnap’s own account, A and B have the same content if and only if the intension of $A \leftrightarrow B$ is necessary. On Chalmers’ account (which we discussed in §5), A and B have the same ideal epistemic content if and only if the primary intension of $A \leftrightarrow B$ is necessary. On our account, A and B have indistinguishable epistemic n -contents if and only if the *epistemic n -intension* of $A \leftrightarrow B$ is necessary. Yet, Jago’s framework does not have the resources to individuate content by appeal to a notion of necessity, since nothing is necessary relative to $W_{\mathcal{J}}$. We thus take it to be an advantage of our account that it allows us to retain the traditional Carnapian idea of understanding semantic content in terms of a kind of necessity.

7 Concluding remarks

In the foregoing, we have outlined a worlds-based theory of semantic content built around a version of Frege’s equipollence principle for content individuation. The semantics identifies the content of a sentence with its epistemic n -intension, which is defined as a function from n -consistent scenarios to extensions. On the proposed picture, two sentences A and B are associated with indistinguishable epistemic n -contents if and only if the epistemic n -intension of $A \leftrightarrow B$ is necessary. We have shown that, for appropriate values of n , non-trivially equivalent sentence pairs such as (1) to (3) get associated with distinguishable epistemic n -contents, whereas trivially equivalent sentence pairs such as (4) to (6) get associated with indistinguishable epistemic n -contents. As such, the proposed semantics offers a joint solution to what

we called the “old” and the “new” problem of hyperintensionality.

While we have focused our efforts on individuating the semantic content of *sentential* expressions, one might naturally wonder whether our semantic framework can be extended to also apply to *subsentential* expressions such as singular terms, predicates, and the like. We are optimistic that the core elements of our semantic framework can be used to develop a generalized version of (Fregean thesis) that applies to sentences and subsentences alike. However, there is a plethora of complicated issues pertaining to how the semantic content of sentences relate to the contents of their subsentential parts.²⁹ In particular, there are tricky issues about how we should populate scenarios with objects and properties when we understand scenarios linguistically as (potentially inconsistent) sets of sentences in some world-making language.³⁰ For this reason, we defer a detailed investigation of the subsentential case to future work.

One might also wonder whether epistemic *n*-intensions may be used to account for various other semantic phenomena concerning how language conveys information and expresses thought content. For instance, we suspect that epistemic *n*-intensions can help us to explain how utterances of mathematical or logical truths can be informative for ordinary persons. On the standard intensional theory of information, a sentence is informative if and only if its intension is contingent.³¹ Consequently, all logical and mathematical truths are deemed uninformative in the intensional framework. By contrast, mathematical and logical truths may well have contingent epistemic *n*-intensions. As such, our framework promises to capture the sense in which mathematical and logical truths can be informative for ordinary persons. Relatedly, epistemic *n*-intensions might also help to explain how deductive reasoning can be informative for ordinary agents. In the intensional framework, $A \rightarrow B$ will be necessary whenever A entails B . Consequently, the intensional framework has the counterintuitive implication that deductive reasoning never carries any information. By contrast, the epis-

²⁹For a survey of questions and problems in this direction, see Szabó (2017).

³⁰For further discussion of these issues, see Jago (2014, §5.6) and Chalmers (2011, §7).

³¹See Bar-Hillel and Carnap (1952) for an early exposition of the intensional account of information.

temic n -intension of $A \rightarrow B$ may well be contingent, even if A entails B . As such, there is reason to hope that the “scandal of deduction”—as Hintikka (1973) called it—can be avoided in our framework.

Although we believe that epistemic n -intensions will have many interesting applications, we are not claiming that our framework can be used to shed light on *all* hyperintensional phenomena that one might want to reason about. For instance, Nolan (2014) surveys a number of interesting and important issues in the area of hyperintensional *metaphysics* that lie beyond the scope of our framework. Moreover, we are not claiming that epistemic n -intensions can be used to shed light on all the *semantic* phenomena that one might want to reason about. Epistemic n -intensions capture a notion of semantic content that is sensitive to the cognitive resources that ordinary speakers have available when they judge and reason semantically. As such, if we are interested in capturing an externalist notion of content—as championed by, e.g., Kripke (1980) and Burge (1979a; 1979b)—epistemic n -content may not be of much use. In particular, since all inference rules in \mathcal{S} are a priori, no sentences in our framework will display the type of necessity that a posteriori necessities such as “water is H_2O ” arguably display. But even if epistemic n -content is not able to play *all* the semantic roles that one might care about, we nonetheless hope to have illustrated its use and value in semantic theorizing.

Let us close by addressing a residual worry that one might have about our semantic framework. The worry centers on the observation that our framework is ‘proof-theoretic’ in the sense that the central notions of (n -consistency) and (n -expansion) are defined in terms of a priori inferability (i.e., inferability in \mathcal{S}). This prompts the worry that our framework, although superficially worlds-based, does not inherit the semantic virtues of a traditional possible-worlds framework, which does not rely on any proof-theoretic structure.

We want to offer two remarks in response to this worry—a dialectical remark, and a substantive one. First, we take one of the central lessons of Bjerring and Schwarz’s impossibility result (discussed in §2) to be that *any* framework that promises a joint solution to the old and the new problem of

hyperintensionality will have to employ some sort of proof-theoretic structure. In particular, as we have seen, this holds of Jago’s hyperintensional framework, which—much like ours—trades on a proof-theoretic relation among worlds. So to the extent that there are problems involved in appealing to proof-theoretic notions within a worlds-based framework, they are not specific to our account, but apply more broadly.³² Second, it seems to us that our framework indeed preserves many interesting features of a traditional possible-worlds framework. In particular, as we showed in §3, epistemic n -intensions play many of the same theoretical roles as classical intensions. Thus, although much more can and should be said about the merits or demerits of appealing to proof-theoretic notions in a semantic context, we find it worthwhile to further develop and apply the proposed semantic framework.

Acknowledgments. Earlier versions of this paper were presented at Massey University, National University of Singapore, Umeå University, and University of Auckland. We thank the audience on those occasions for valuable feedback. A special thanks to two anonymous reviewers for *Synthese*, and to Andreas Stokke, Ben Blumson, John Matthewson, Lars Bo Gundersen, Weng Hong Tang, and Wolfgang Schwarz for very helpful comments and criticism on earlier versions of this paper.

References

- Barwise, J. and J. Perry (1983). *Situations and Attitudes*. University of Chicago Press.
- Berto, F. (2013). “Impossible Worlds”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2013. Metaphysics Research Lab, Stanford University.
- Bjerring, J. C. (2010). “Non-Ideal Epistemic Spaces”. PhD thesis. Australian National University.

³²For further considerations in this direction, see Bjerring and Schwarz (2017, §§5-6). See also Chalmers (2011), Bjerring (2010), and Jago (2009) for discussions of how proof-theoretic structure might help to construct a ‘non-ideal epistemic space’ that is sensitive to the limited cognitive resources of non-ideal epistemic agents.

- Bjerring, J. C. (2013). “Impossible Worlds and Logical Omniscience: An Impossibility Result”. In: *Synthese* 190.13, pp. 2505–2524.
- Bjerring, J. C. and W. Schwarz (2017). “Granularity Problems”. In: *Philosophical Quarterly* 67.266, pp. 22–37.
- Bjerring, J. C. and M. Skipper (forthcoming). “A dynamic solution to the problem of logical omniscience”. In: *Journal of Philosophical Logic*.
- Burge, T. (1979a). “Individualism and the Mental”. In: *Midwest Studies in Philosophy* 4.1, pp. 73–122.
- (1979b). “Sinning Against Frege”. In: *Philosophical Review* 88.3, pp. 398–432.
- Carnap, R. (1947). *Meaning and Necessity*. University of Chicago Press.
- Carnap, R. and Y. Bar-Hillel (1952). *An Outline of a Theory of Semantic Information*. Tech. rep. 247. MIT.
- Chalmers, D. (2002). “On Sense and Intension”. In: *Philosophical Perspectives* 16.s16, pp. 135–82.
- (2006). “The Foundations of Two-Dimensional Semantics”. In: *Two-Dimensional Semantics: Foundations and Applications*. Ed. by Manuel Garcia-Carpintero and Josep Macia. Oxford University Press, pp. 55–140.
- (2011). “The Nature of Epistemic Space”. In: *Epistemic Modality*. Ed. by Andy Egan and Brian Weatherson. Oxford University Press.
- Drapkin, J. and D. Perlis (1986). “A Preliminary Excursion into Step-Logics”. In: *Proceedings of the SIGART International Symposium on Methodologies for Intelligent Systems*. Ed. by C. Ghidini, P. Giodini, and W. van der Hoek, pp. 262–269.
- Elliott, E. (2017). “Impossible worlds and partial belief”. In: *Synthese* (online first).
- Faroldi, F. (2017). “Co-Hyperintensionality”. In: *Ratio* 30.3, pp. 270–287.
- Frege, G. (1892). “On Sense and Reference [Über Sinn und Bedeutung]”. In: *Zeitschrift für Philosophie und philosophische Kritik* 100, pp. 25–50.
- (1969/1979). *Nachgelassene Schriften*. Hamburg: Felix Meiner (P. Long and R. White, Trans. Posthumous writings). Chicago: University of Chicago Press (1979).

- Frege, G. (1984). *Collected Papers on Mathematics, Logic, and Philosophy*. Oxford: Basil Blackwell.
- Glanzberg, M. (2015). “Logical Consequence and Natural Language”. In: *Foundations of Logical Consequence*. Ed. by Colin Caret and Ole Hjortland. Oxford University Press, pp. 71–120.
- Hintikka, J. (1973). *Logic, Language-Games and Information: Kantian Themes in the Philosophy of Logic*. Oxford, Clarendon Press.
- (1975). “Impossible Possible Worlds Vindicated”. In: *Journal of Philosophical Logic* 4, pp. 475–484.
- Jago, M. (2009). “Logical Information and Epistemic Space”. In: *Synthese* 167.2, pp. 327–41.
- (2013b). “The Problem of Rational Knowledge”. In: *Erkenntnis* 6, pp. 1–18.
- (2014). *The Impossible: An Essay on Hyperintensionality*. Oxford University Press.
- King, J. (2016). “Structured Propositions”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2016.
- Kripke, S. (1980). *Naming and Necessity*. Harvard University Press.
- Leitgeb, H. (ms.). “HYPE: A System of Hyperintensional Logic”.
- Lewis, D. (2004). “Letters to Priest and Beall”. In: *The Law of Non-Contradiction*. Oxford University Press, pp. 176–177.
- Nolan, D. (2014). “Hyperintensional Metaphysics”. In: *Philosophical Studies* 171.1, pp. 149–160.
- (forthcoming). “Impossibility and Impossible Worlds”. In: *The Routledge Handbook of Modality*. Ed. by Otavio Bueno and Scott Shalkowski. New York, USA: Routledge Press.
- Putnam, H. (1975). “The Meaning of ‘Meaning’”. In: *Minnesota Studies in the Philosophy of Science* 7, pp. 131–193.
- Rasmussen, M. Skipper (2015). “Dynamic Epistemic Logic and Logical Omniscience”. In: *Logic and Logical Philosophy* 24, pp. 377–399.
- Russell, B. (2017). “A Priori Justification and Knowledge”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. Zalta. Summer 2017. Metaphysics Research Lab, Stanford University.

- Smets, S. and A. Solaki (2018). “The Effort of Reasoning: Modelling the Inference Steps of Boundedly Rational Agents”. In: *Logic, Language, Information, and Computation - Proceedings of 25th International Workshop*, pp. 307–24.
- Stalnaker, R. (1984). *Inquiry*. Cambridge University Press.
- (2012). “Intellectualism and the Objects of Knowledge”. In: *Philosophy and Phenomenological Research* 85.3, pp. 754–61.
- Stanley, J. (2010). “"Assertion" and Intentionality”. In: *Philosophical Studies* 151.1, pp. 87–113.
- (2011). *Know How*. Oxford University Press.
- Szabó, Z. (2017). “Compositionality”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. Zalta. Summer 2017. Metaphysics Research Lab, Stanford University.
- Zimmermann, T. (1999). “Meaning Postulates and the Model-Theoretic Approach to Natural Language Semantics”. In: *Linguistics and Philosophy* 22.5, pp. 529–561.