Hyperintensional semantics: a Fregean approach

Jens Christian Bjerring
Mattias Skipper Rasmussen

Draft, December 2016

Abstract

In this paper, we present a new theory of semantic content that inherits two central aspects of the Fregean notion of *sense*. First, it inherits the *hyperintensional* aspect of Fregean senses. Senses are hyperintensional because some necessarily equivalent sentences have distinct senses. Second, it inherits the *non-trivial* aspect of Fregean senses. Senses are non-trivial because some necessarily equivalent sentences have the same sense. So, while cut finer than necessary equivalence, sense is not an "anything goes" view on content: there are substantial constraints on how finely individuated Fregean senses should be. While traditional worlds-based theories of semantic content have been unable to accommodate content that—much like Fregean senses—is both hyperintensional and non-trivial, we show that our semantics successfully models such content in virtue of validating Frege's so-called *equipollence principle* for sense individuation. We also argue that our semantic framework can be used to solve a number of outstanding philosophical problems that call out for a hyperintensional, yet non-trivial analysis, and illustrate by providing new solutions to *the scandal of deduction* and *the problem of logical omniscience*.

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

1 Introduction

In this paper, we present a new theory of semantic content that inherits two central aspects of the Fregean notion of *sense*. First, it inherits the *hyperintensional* aspect of Fregean senses. Senses are hyperintensional because some necessarily equivalent sentences have distinct senses. Consider the following pairs of equivalent 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. $(A \wedge (\neg(\neg A \vee B))) \vee (A \wedge B)$.

These sentence pairs differ in sense because it can be cognitively significant to be told that they are equivalent: it requires substantial cognitive work to establish the equivalences, and they can provide us with novel information about the world. Consequently, these sentences can also play different roles in an agent's cognitive life: I may well believe A but fail to believe $(A \wedge (\neg(\neg A \vee B))) \vee (A \wedge B)$, or intend to communicate that Hesperus is Phosphorus, but not the trivial claim that Hesperus is self-identical.

Second, our semantics will inherit the *non-trivial* aspect of Fregean senses. Senses are non-trivial because some necessarily equivalent sentences have the same sense. Consider the following pairs of equivalent sentences:[1]

(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. $A \wedge A$.

[1]The pairs (5) and (6) are extracted from Frege (1984, pp. 390-406).

These sentence pairs agree in sense because it is cognitively insignificant to be told that they are equivalent: it requires only trivial cognitive work to establish the equivalences, and they seem to provide us with no new information about the world. As such, the sentences will also play very similar roles in an agent's cognitive life: anyone who believes that x equals y will also believe that y equals x, and anyone who intends to communicate that the number of primes is infinite will also intend to convey that there are infinitely many primes. So, while cut finer than necessary equivalence, sense is not an "anything goes" view on content: there are substantial constraints on how finely individuated senses should be.

We will use Frege's so-called *equipollence principle* for sense individuation to capture these intuitions about the hyperintensional, yet non-trivial nature of semantic content. According to this principle, two sentences A and B agree in sense just in case 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 1979, p. 197). So essentially, the equipollence principle individuates semantic content by reference to a relation of being *immediately recognizable as equivalent*:

> **(Fregean thesis$_1$)** Two sentences A and B have the same content if and only if A and B are immediately recognizable as equivalent.

While the sentence pairs in (4) to (6) will be immediately recognizable as equivalent for most ordinary agents, those in (1) to (3) will not. For instance, while it is immediately clear for an ordinary rational agent that A is equivalent to $A \wedge A$, it requires significant cognitive work to establish that A is equivalent to $(A \wedge (\neg(\neg A \vee B))) \vee (A \wedge B)$. So (Fregean thesis$_1$) captures a notion of semantic content that is hyperintensional and yet non-trivial: while non-trivially equivalent sentence pairs such as (1) to (3) are associated with distinct contents, trivially equivalent sentence pairs such as (4) to (6) are associated with the same content.

Why should we care about a *hyperintensional* semantics? Because it can help us overcome a central problem for standard intensional semantics:

namely that it only yields coarse-grained distinctions among contents. In the intensional framework, semantic content is represented by intensions, which are defined as functions from possible worlds to extensions. And since any two necessarily equivalent sentences are co-extensional at all possible worlds, such sentences are associated with identical intensions, and hence with identical contents. But if we take the Fregean thoughts about content individuation seriously, as we do, there are many pairs of equivalent sentences that should be associated with distinct contents. So intensional content is overly coarse-grained for our purposes.

In a hyperintensional framework, semantic content is instead represented by hyperintensions, which are defined as functions from impossible worlds to extensions. Intuitively, while possible worlds represent ways the world *could* be, *impossible* worlds represent ways the world could *not* be: worlds where the metaphysically, analytically, mathematically, or logically impossible may happen. Since necessarily equivalent sentences need not be co-extensional at impossible worlds, they can be associated with distinct hyperintensions, and hence with distinct contents.

Why, then, should we care about a *non-trivial* hyperintensional semantics? Because it can help us overcome a central risk that hyperintensional accounts of content face: namely that of drawing distinctions among contents where, intuitively, there should be none. Those who adopt an impossible worlds framework often take a liberal stance on the logical structure of impossible worlds. In particular, it is often assumed that *any* two necessarily equivalent sentences differ in truth-value at some impossible world—see Nolan (1997), Priest (2005), Vander Laan (1997), Zalta (1997), and others. Given this assumption, it follows that any two equivalent sentences are associated with distinct hyperintensions, and hence with distinct contents. But if we take the Fregean thoughts about content individuation seriously, as we do, there are many pairs of equivalent sentences that should agree in content. So, on pain of making hyperintensional content too fine-grained, impossible worlds should display substantial logical structure.

In what follows, we will use Frege's equipollence principle to develop a worlds-based semantics that can steer clear of notions of content that are either implausibly coarse-grained or implausibly fine-grained. More specifically, our semantics will allow us to reason about a whole *spectrum* of notions of content that ranges from extremely fine-grained—and potentially trivial—content to content that is individuated up to a priori equivalence. The resulting semantic flexibility, as we shall see, flows directly from the equipollence principle and is desirable because it is unlikely that a single, intermediately-grained notion of content applies across all hyperintensional contexts.

While we will adopt a worlds-based approach to hyperintensional semantics, there are alternatives. First, there are *skeptics* about hyperintensionality such as Stalnaker (1984) who argues that seemingly hyperintensional distinctions can be explained away meta-linguistically. Second, there are *structuralists* about hyperintensionality who aim to capture hyperintensional distinctions by understanding propositions as structured entities consisting of, say, objects, properties, functions, and relations.[2] While we believe there are serious problems with both the sceptic and the structuralist approach to hyperintensionality, this is not the place to air these misgivings.[3]

Although there is an obvious Fregean flavor to our project, our aim is not to give an exegetical account of Frege's conception of sense, nor to defend a broadly Fregean approach to semantics.[4] Rather, our aim is to *use* Frege's equipollence principle to reason about a notion of semantic content that is both hyperintensional and non-trivial. And while such a notion of "intermediately-grained" content can do important work in different areas of philosophy, as we will illustrate in §5, we are not claiming that it can do *all* the work that one might ask of a semantic theory. The equipollence principle offers a *cognitive* or *epistemic* individuation of content, and, as we will discuss further in §6, such an individuation principle might not be useful for reasoning about semantic phenomena involving, say, consequence relations

[2] See King (2016) for an overview of various structuralist approaches to semantic content.

[3] For a detailed critique of these alternative approaches to hyperintensionality, see Stanley (2010) and Jago (2014, ch. 2-3).

[4] For an exegetical discussion of Frege's equipollence principle for sense individuation, see Penco (2003) and Schellenberg (2012).

in natural languages or communication between speakers. But we take it that there is no conflict here. There are many things that we might want a semantic theory to do. On our picture, the Fregean approach to semantics earns its keep by the work it can do in helping us reason about how language conveys information and expresses thought contents.

With these remarks in place, here is the plan for the rest of the paper. In §2 we refine our interpretation of Frege's equipollence principle. In §3 we use the refined equipollence principle to define a special class of non-trivial hyperintensions. In §4 we use these hyperintensions to formalize Frege's equipollence principle, and argue that the resulting semantics yields the correct verdicts in paradigmatic Fregean cases such as (1) to (6). We also criticize a related hyperintensional account of Fregean content that has recently been proposed by Mark Jago (2014). Finally, in §5, we argue that our semantic framework can be used to solve a number of open philosophical problems that call out for a hyperintensional, yet non-trivial analysis. We illustrate by providing new solutions to *the scandal of deduction* in philosphy of information and *the problem of logical omniscience* in formal epistemology.

2 The Fregean thesis

We begin by refining various elements of (Fregean thesis$_1$). First, note that (Fregean thesis$_1$) imposes a *cognitive* or *epistemic* constraint on content individuation. According to (Fregean thesis$_1$), for two sentences to agree in content, it is not enough for them to be equivalent—they have to be so in a sufficiently "immediate" or "trivial" manner. 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 very complex mathematical and logical equivalences can have the same content, according to (Fregean thesis$_1$), whereas this is not so if we focus on people as they are most. So the cuts in semantic content that (Fregean thesis$_1$) provides varies with agents' cognitive resources. The Fregean approach to semantic content hence differs from the kind of externalist approach we find in Kripke (1980), Putnam (1975), and Burge (1979a; 1979b). Roughly, on their view, the semantic content of a linguistic expression is determined by the referential relations it bears to various external or mind-independent entities. And since such referential relations remain stable even when sense varies, semantic content does not vary with cognitive resources on their view. To reflect the epistemic or cognitive individuation of semantic content in Frege's equipollence principle, we will henceforth talk about the *epistemic content*—or, if you prefer, the *cognitive content*—of linguistic expressions.

Note, however, that although content individuation on the Fregean picture varies with cognitive resources, it does *not* vary with whatever empirical information an agent happens to have. Strictly speaking, what is immediately recognizable does depend on the empirical information that an agent has available. For instance, if an agent has a lot of information about astronomy, it may be immediately recognizable for her that 'Hesperus is Phosphorus' and 'Hesperus is Hesperus' are equivalent. Yet, since Frege's equipollence principle explicitly quantifies over *all* agents—regardless of their empirical information—content individuation is insensitive to arbitrary stocks of empirical information (we return to this point in §4.1).

Second, the relation of being immediately recognizable as equivalent is a *intransitive* relation. To see why, consider two sentences S_1 and S_n that are equivalent, but only highly non-trivially so. In particular, assume that it takes a large number n of applications of deductive rules to infer S_n from S_1, and vice versa. For ordinary rational agents like you and me, S_1 and S_n will not count as immediately recognizable as equivalent because it is highly cognitively demanding to establish their equivalence. Yet, since any one application of a simple deductive rule leading from S_i to S_{i+1} is immediate or trivial, each pair of adjacent sentences S_i and S_{i+1} *will* be immediately recognizable as equivalent for ordinary rational agents. So the relation of being immediately recognizable as equivalent is intransitive: A need not be immediately recognizable as equivalent to C just because A is immediately recognizable as equivalent to B, and B is immediately recognizable as equivalent to C.

But then (Fregean thesis$_1$) cannot be the right interpretation of Frege's equipollence principle. For according to (Fregean thesis$_1$), two sentences have the *same* content whenever they are immediately recognizable as equivalent, and content sameness is most naturally understood as an identity relation, which is transitive. So we need to replace 'content sameness' on the left-hand side of (Fregean thesis$_1$) with an intransitive substitute that is faithful to the epistemic aspect of Frege's equipollence principle. Stipulatively, we will refer to this intransitive relation as *epistemic content indistinguishability*:

> **(Fregean thesis$_2$)** Two sentences A and B have indistinguishable epistemic contents if and only if A and B are immediately recognizable as equivalent.

We can think of the intransitivity of epistemic content indistinguishability as mirroring the intransitivity of perceptual content indistinguishability. Consider a graded color spectrum that ranges from black to white. While each pair of adjacent lines in the color spectrum are indistinguishable to the human eye, a line towards the black end of the spectrum *will* be distinguishable from a line towards the white end of the spectrum. So the relation of having indistinguishable perceptual content is intransitive. Similarly, while each pair of adjacent sentences S_i and S_{i+1} in the complex deductive chain above will play indistinguishable cognitive roles for ordinary rational agents, S_1 and S_n will play distinguishable such roles. So the relation of having indistinguishable epistemic content is intransitive.

The intransitivity observation is no mere terminological quibble: it shows importantly that we cannot model what it means for two sentences to have indistinguishable epistemic contents by determining whether their individual contents are identical. To appreciate the significance of this point, suppose we work in classical intensional semantics and represent content in terms of intensions. We might then hold that two sentences have indistinguishable epistemic contents just in case they have identical intensions. Yet, this would wrongly treat content indistinguishability as a transitive relation. The point obviously generalizes: regardless of whether semantic content is represented by intensions, hyperintensions, structured propositions, mental states, or the like, we cannot determine whether two sentences have indistinguishable epistemic contents by looking at whether their individual contents are identical.[5]

The task, then, is to find an intransitive relation that can capture what it means for two sentences to be immediately recognizable as equivalent, and hence capture what it means for two sentences to have indistinguishable epistemic contents. Generally, what is immediately recognizable as equivalent is closely tied 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 because it is immediately a priori inferable that they are equivalent. For instance, A and $A \wedge A$ are immediately recognizable as equivalent because the inference from A to $A \wedge A$, and vice versa, is a priori and immediate. Similarly, $x = y$ and $y = x$ are immediately recognizable as equivalent because the inference from $x = y$ to $y = x$, and vice versa, is a priori and immediate. More generally, what is immediately a priori inferable depends—just like what is immediately recognizable as equivalent—on the cognitive resources that agents have available for a priori reasoning. While it may be immediate or trivial for a skilled mathematician to establish that the interior angles of a hexagon in Euclidian space must sum up to 4π, this line of reasoning may well be non-trivial for you and me.

To capture how immediate a priori inferability depends on agents' (potentially) limited cognitive resources, we need a measure of the complexity of different chains of a priori reasoning. For present purposes, little will hinge on whether we specify this measure by reference to time consumption, neural firings in the brain, or some third quantity. In principle, any quantitative measure can do the job here. As a simple model, however, we will measure an agent's cognitive resources in terms of the number n of steps of a priori reasoning that the agent can immediately or trivially perform. The higher the value of n, the more are the cognitive resources that the agent has available for immediate a priori reasoning. We will think of a priori reasoning as a rule-bound process whereby agents reason, one step at the time, by applying rules from a designated set $\mathcal{S}$ of a priori inference rules.[6] We can understand

[5] For a general discussion of the intransitivity of semantic content, see Bjerring & Schwarz (forthcoming).

[6] For similar ways of modelling resource-bounded reasoning, see Drapkin & Perlis (1986), Jago (2014), Bjerring (2010; 2013; 2014), and others.

$\mathcal{S}$ 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. While we are familiar with thinking about logical reasoning in terms of rules such as modus ponens and disjunctive syllogism, it is less clear how to encode mathematical and conceptual reasoning in $\mathcal{S}$, as we shall discuss further in §4. For now it suffices to note that all rules in $\mathcal{S}$—whether logical or non-logical—should be *a priori*. So, for instance, there will be no rules in $\mathcal{S}$ that allow agents to move inferentially from sentences containing terms such as 'Hesperus' to sentences containing terms such as 'Phosphorus.' For it requires significant astronomical information to move from facts about Hesperus to facts about Phosphorus.

We can then encode immediate a priori reasoning as n-step reasoning in $\mathcal{S}$ and say that B is immediately a priori inferable from A just in case B can be inferred from A by at most n applications of the rules in $\mathcal{S}$—that is, just in case B is *n-step inferable in $\mathcal{S}$* from A. In particular, if A is n-step inferable in $\mathcal{S}$ from the empty set, we will simply say that A is n-step inferable in $\mathcal{S}$. Also, if B follows from A by some number of applications of the rules in $\mathcal{S}$, we will say that B is *in principle* inferable in $\mathcal{S}$ from A. For ease of exposition, we will throughout assume that nothing is 0-step inferable in $\mathcal{S}$, except for the trivial inference from A to A.

Since we understand what is immediately recognizable as equivalent in terms of what is n-step inferable in $\mathcal{S}$, we will henceforth refer to the relevant notion of content as *epistemic n-content*. We can then refine our interpretation of Frege's equipollence principle as follows:

> **(Fregean thesis$_3$)** 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.

It is easily seen that (Fregean thesis$_3$) individuates epistemic n-content in an intransitive way: C need not be n-step inferable from A just because B is n-step inferable from A, and C is n-step inferable from B. For instance, if we stipulate that $\mathcal{S}$ 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, as intended, the notion of n-step inferability in $\mathcal{S}$ provides us with an intransitive relation that captures the epistemic aspect of Frege's equipollence principle.

Importantly, (Fregean thesis$_3$) gives rise to a whole *spectrum* of notions of content with different levels of granularity. In the limit where $n = 0$, (Fregean thesis$_3$) individuates epistemic content in an extremely fine-grained manner: any two distinct sentences will have distinguishable epistemic n-contents (assuming, as we do, that nothing is 0-step inferable in $\mathcal{S}$). In the opposite limit where n goes towards infinity, (Fregean thesis$_3$) individuates epistemic content in a coarse-grained manner: any two a priori equivalent sentences will have indistinguishable epistemic n-contents since, by stipulation, all equivalences are in principle inferable in $\mathcal{S}$. In between these extremes, we find a spectrum of notions of content that—much like Fregean senses—are hyperintensional, yet non-trivial: some, but not all priori equivalent sentences will have indistingushable epistemic n-contents since, for intermediate values of n, some, but not all a priori equivalences are n-step inferable in $\mathcal{S}$.

With these refinements of Frege's equipollence principle on the table, we can now develop a worlds-based semantics that validates (Fregean thesis$_3$).

3 Epistemic n-intensions

The guiding idea behind our semantics is to understand n-step inferability in $\mathcal{S}$ as a kind of (epistemic) necessity that is sensitive to the cognitive resources that agents have available for a priori reasoning. More specifically, we will say that a sentence A is necessary if and only if A is n-step inferable in $\mathcal{S}$. Since only a priori truths are inferable in $\mathcal{S}$, only a priori truths can be necessary, according to this definition, and *which* a priori truths are necessary will depend on the value of n. When $n = 0$, no a priori truths are necessary since no such truths are 0-step inferable in $\mathcal{S}$. When n goes towards infinity, all a priori truths are necessary since all such truths are in principle inferable in $\mathcal{S}$. For intermediate values of n, some, but not all a priori truths are necessary.

Treating n-step inferability in $\mathcal{S}$ as a kind of necessity will allow us to develop a semantic analysis of (Fregean thesis$_3$) that closely mirrors the following familiar content 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.

We will let our hyperintensional notion of necessity play largely the same role in our analysis of (Fregean thesis$_3$) that the standard notion of necessity plays in (Intensional content). More precisely, we will define a class of non-trivial hyperintensions—so-called *epistemic n-intensions*—that will eventually allow us to say that A and B have indistinguishable epistemic n-contents just in case the epistemic n-intension of $A \leftrightarrow B$ is necessary.

To spell out the details of this proposal, we first need to construct a space of scenarios at which we can evaluate sentences for truth and falsity. Informally, we will think of scenarios as representations of ways the world might be for all an agent can tell a priori. Since some agents have limited resources available for a priori reasoning, some scenarios will give an a priori inconsistent representation of a way the world might be. Such scenarios are *a priori impossible*: they look possible to agents with limited capacities for a priori reasoning, but are nonetheless a priori inconsistent. Since some (idealized) agents have unlimited resources available for a priori reasoning, some scenarios will give an a priori consistent representation of a way the world might be. Such scenarios are *a priori possible*: no amount of a priori reasoning will reveal these scenarios as inconsistent.

We will construct both possible and impossible scenarios linguistically as sets of sentence types in a language $\mathcal{L}$ that is rich enough to represent every possible and impossible way the world might be for agents with (potentially) limited resources available for a priori reasoning. Obviously, much can be said about the specifics of such a world-making language, but for present purposes it suffices to specify a few general requirements that $\mathcal{L}$ should meet. 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 define hyperintensions that are supposed to mirror Fregean content. Second, $\mathcal{L}$ should contain a sufficiently rich vocabulary that allows us to express the kinds of sentences that we find in logic, mathematics, science, and so on. Third, $\mathcal{L}$ should display no context-dependent features. Had we constructed scenarios in a context-dependent language like English, scenarios need not give determinate representations of ways the world might be. For instance, if scenarios include demonstrative sentences like '*this* is blue,' it would be unclear what such scenarios represent, and hence unclear what would be true and false at such scenarios. But aside from such general requirements, we need not make any assumptions about the specific nature of $\mathcal{L}$.

Since we want to evaluate sentence tokens in English at scenarios constructed in $\mathcal{L}$, we need a way to relate tokens in English to sentence types in $\mathcal{L}$. To this end, we can stipulate a translation function that maps tokens in English to types in $\mathcal{L}$ in a way that respects the fine grained content of types in $\mathcal{L}$. In particular, we can ensure that sentence pairs such as (1) to (6) always have distinct contents in $\mathcal{L}$ by requiring that the translation function maps any two tokens of orthographically distinct sentence types in English to distinct sentence types in $\mathcal{L}$. But again, aside from such general requirements, we need not make any specific assumptions about the translation function.

Given the linguistic construction of scenarios as sets of sentences in $\mathcal{L}$, we can go on to specify what it means for sentence tokens $A, B, C, \ldots$ in English to be true at scenarios. For ease of exposition, we shall ignore the relation between English and the world-making language $\mathcal{L}$ and pretend that scenarios correspond directly 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. Let V be an evaluation function from sentences and scenarios to truth-values. We can then define:

(Verification) $V(A, w) = T$ if and only if $A \in w$.[7]

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

[7] On the official story on which only sentence types in $\mathcal{L}$ can constitute scenarios, we should say that $V(A, w) = T$ if and only if $A^+ \in w$, where A^+ is the sentence type in $\mathcal{L}$ that translates the sentence token A in English.

If a scenario neither verifies nor falsifies A, we will say that A is *indeterminate* at the scenario. If a scenario contains one or more indeterminacies, we will say that the scenario is *partial* or *gappy*.

While any scenario corresponds to a set of sentences, not any set of sentences should correspond to a scenario. For scenarios should represent ways the world might be for agents with (potentially) limited resources available for a priori reasoning, and there are constraints on what the world might be like for such agents. Intuitively, scenarios that can immediately or trivially be shown to be inconsistent should not count as ways the world might be for such agents.[8] More precisely, since we focus on agents who can trivially or immediately perform n steps of a priori reasoning, we can restrict our attention to scenarios from which no hidden contradiction can be revealed within n steps of a priori reasoning. That is, we can restrict our attention to scenarios that are *n-consistent* (where '$\Gamma \vdash^n_{\mathcal{S}} \Gamma'$' means that the set of sentences Γ' can be inferred from the set Γ by n applications of the rules in $\mathcal{S}$):

> **(n-consistency)** A scenario w is n-consistent if and only if $w \nvdash^n_{\mathcal{S}} \perp$, where $\perp$ is an arbitrary contradiction of the form $\{A, \neg A\}$.

To illustrate, suppose $\mathcal{S}$ contains just modus ponens, and let $w = \{A, A \rightarrow B, \neg B\}$. Since it takes one application of modus ponens to derive an explicit contradiction from w, w is 0-consistent but 1-inconsistent. By contrast, the scenario $v = \{A, A \rightarrow B, B \rightarrow C, \neg C\}$ is 1-consistent, but 2-inconsistent since it takes two applications of modus ponens to derive an explicit contradiction from v. The scenario $v' = \{A \vee B\}$ is consistent *simpliciter* since it is not possible to derive a contradiction from v' in any number of steps using the rules in $\mathcal{S}$.

We can now define W_n as 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 $\{A, \neg A\}$. So while scenarios in W_n may

[8]This 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) for similar ideas.

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 scenarios in W_n. When $n = 0$, W_n is highly permissive: it only excludes scenarios that verify explicit contradictions. When n goes towards infinity, W_n is highly restrictive: it only permits scenarios that are a priori consistent *simpliciter*. For intermediate values of n, scenarios in W_n will display different intermediate degrees of consistency. So we get a whole spectrum of more or less permissive modal spaces that can be ordered in a nested sequence:

$$W_0 \supseteq W_1 \supseteq W_2 \supseteq \cdots$$

These subset relations say that any j-consistent scenario will be i-consistent, for $i \leq j$, and, conversely, that any i-inconsistent scenario will be j-inconsistent. This holds because any contradiction that cannot be inferred within n steps also cannot be inferred within fewer than n steps, and, conversely, because any contradiction that can be inferred within n steps can also be inferred within more than n steps.

While (n-consistency) guarantees that no scenario in W_n will be trivially inconsistent, it does not by itself give us the required tools to treat n-step inferability in $\mathcal{S}$ as a kind of necessity. For on the usual way of understanding necessities in worlds-based frameworks, a sentence is necessary just in case it is true across a certain class of scenarios. But since (n-consistency) does not constrain the number of indeterminacies or truth-value gaps that scenarios in W_n can contain, no sentence will be true at all scenarios in W_n, regardless of the value of n. So to ensure that our framework yields an appropriate class of necessities, we need a tool that eliminates truth-value gaps from scenarios in W_n:

> **(n-expansion)** For any scenario $w \in W_n$, the n-expansion of w is the scenario $w^n = \{A \mid w \vdash^n_{\mathcal{S}} A\}$.

Intuitively, (n-expansion) can help us "fill up" partial scenarios in W_n: the n-expansion of w contains every sentence contained in w *plus* every sentence that can be inferred from w within n steps of reasoning in $\mathcal{S}$. To illustrate, let

$w = \{\neg\neg(A \wedge B)\}$, and suppose that $\mathcal{S}$ contains just conjunction elimination and double negation elimination. The 1-expansion of w is then $w^1 = \{\neg\neg(A \wedge B), A \wedge B\}$, whereas the 2-expansion of w is $w^2 = \{\neg\neg(A \wedge B), A \wedge B, A, B\}$. Given that nothing is 0-step inferable in $\mathcal{S}$, the 0-expansion of w is identical to w itself. Note also that since scenarios in W_n are n-consistent, the n-expansion of a scenario in W_n will never verify an explicit contradiction.[9]

We can then define the *epistemic n-intension* of a sentence A 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 $V(A, w^n) = F$.

If the epistemic n-intension of A is neither true nor false at w, we will say that it is *indeterminate* at w. We can think of the epistemic n-intension of a sentence as specifying the conditions under which the sentence is true at scenarios that represent ways-the-world-might-be for agents with no empirical information and (potentially) limited resources available for a priori reasoning. By varying the value of n, we get a whole spectrum of epistemic n-intensions. When $n = 0$, epistemic n-intensions will be extremely ill-behaved: they are only guaranteed not to verify and falsify the same sentence at the same scenario. When n approaches infinity, epistemic n-intensions will be highly well-behaved: they will deliver truth-values that fully match what can in principle be established a priori about the world. For intermediate values of n, we find a spectrum of epistemic n-intensions that display different intermediate degrees of conformance with what can in principle be established a priori about the world.

We can now define what it means for the epistemic n-intension of a sentence 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.

[9] Note that (n-expansion) does not amount to a closure constraint on scenarios in W_n: it does not say that scenarios in W_n verify all n-step consequences of what they verify. This is important because it can be shown that any scenario that is closed under n-step consequence is deductively closed *simpliciter*. For further discussion of this result, see Bjerring (2010; 2013; 2014) and Jago (2014).

(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 it is *contingent*. As in other frameworks that allow for truth-value gaps, we cannot interdefine possibility and necessity since the epistemic n-intension of A need not be true at a scenario w just because it is not false at w, nor need it be false at w just because it is not true at w.

Epistemic n-intensions are in many respects comparable to classical intensions. 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 W_n into those scenarios whose n-expansions verify a given sentence, and those whose n-expansions do not. Second, classical intensions can be used to give set-theoretic representations of the meanings of various logical connectives. For instance, the meaning of '$\wedge$' and '$\vee$' can be expressed as follows, where $|A|$ is the set of possible worlds that verify A:

$$(\wedge) \quad \begin{array}{l} |A \wedge B| \subseteq |A| \cap |B| \\ |A| \cap |B| \subseteq |A \wedge B| \end{array} \qquad (\vee) \quad \begin{array}{l} |A \vee B| \subseteq |A| \cup |B| \\ |A| \cup |B| \subseteq |A \vee B| \end{array}$$

It is well-known that we have to give up these classical dependencies once we work in an impossible worlds framework.[10] Nonetheless, we can still use epistemic n-intensions to give informative set-theoretic representations of the roles that various logical connectives play in cognition and reasoning. Let $|A|^n$ be the set of scenarios in W_n at which the epistemic n-intension of A is true, and assume that $\mathcal{S}$ contains the standard introduction and elimination rules for conjunction and disjunction. We can then give the following set-theoretic representations of the meanings of '$\wedge$' and '$\vee$':

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

[10] For details, see Bjerring & Schwarz (forthcoming).

For the upper subset relation involving conjunction, suppose that $w \in |A \wedge B|^n$. We must show that $w \in |A|^{n+1} \cap |B|^{n+1}$. By (Epistemic n-intension) and (Verification), $A \wedge B \in w^n$. And since $\mathcal{S}$ contains conjunction elimination, $w^n \vdash^1_{\mathcal{S}} A$ and $w^n \vdash^1_{\mathcal{S}} B$, which means that $w \vdash^{n+1}_{\mathcal{S}} A$ and $w \vdash^{n+1}_{\mathcal{S}} B$. So, by ($n$-expansion), $A \in w^{n+1}$ and $B \in w^{n+1}$. By (Epistemic n-intension) and (Verification), then, the epistemic $n+1$-intension of A is true at w, and similarly for B. So $w \in |A|^{n+1}$ and $w \in |B|^{n+1}$, in which case $w \in |A|^{n+1} \cap |B|^{n+1}$. The other subset relations can be established in similar ways.

The subset relations above show that although our semantics is not truth-functional in the classical sense, it nevertheless allows us to establish informative dependencies between the epistemic n-intensions of complex sentences and those of their parts. For instance, while the epistemic n-intension of $A \wedge B$ need not be true just because the epistemic n-intensions of A and B are both true, we know that the epistemic $(n+1)$-intension of $A \wedge B$ *must* be true. And while the epistemic n-intension of A need not be true just because the epistemic n-intension of $A \wedge B$ is true, we know that the epistemic $(n+1)$-intension of A *must* be true. So epistemic n-intensions still behave in ways that track the roles that various logical connectives play in cognition and reasoning.

4 Epistemic *n*-content and Fregean content

With the semantic framework in place, we can now represent the epistemic n-content of a sentence by its associated epistemic n-intension, and give our final analysis of Frege's equipollence principle:

> **(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 intended, (Fregean thesis) individuates epistemic n-content in an intransitive way: the epistemic n-intension of $A \leftrightarrow C$ need not be necessary just because the epistemic n-intensions of $A \leftrightarrow B$ and $B \leftrightarrow C$ are both necessary. For instance, while the epistemic 1-intensions of $A \leftrightarrow (A \wedge A)$ and

$(A \wedge A) \leftrightarrow (A \wedge A) \wedge (A \wedge A)$ are both necessary, the epistemic 1-intension of $A \leftrightarrow (A \wedge A) \wedge (A \wedge A)$ is not.

The following result guarantees that (Fregean thesis) succesfully validates (Fregean thesis$_3$):

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

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

As a special case of Theorem 1, the epistemic n-intension of $A \leftrightarrow B$ is necessary just in case $A \leftrightarrow B$ is n-step inferable in $\mathcal{S}$. So (Fregean thesis) and (Fregean thesis$_3$) are extensionally equivalent: they always yield the same semantic verdicts.

By varying the value of n, (Fregean thesis) can be used to reason about a whole spectrum of notions of epistemic content with different levels of granularity. When $n = 0$, (Fregean thesis) individuates content in an extremely fine-grained manner: any two sentences, whether a priori equivalent or not, have distinguishable epistemic n-contents. When n goes towards infinity, (Fregean thesis) individuates content up to a priori equivalence: any two a priori equivalent sentences have indistinguishable epistemic n-contents. For intermediate values of n, (Fregean thesis) captures a range of notions of content that—much like Fregean senses—are hyperintensional, yet non-trivial: some, but not all a priori equivalent sentences have indistinguishable epistemic n-contents.

There are a number of interesting observations concerning the limit case where n approaches infinity. First, while the resulting notion of content is individuated up to a priori equivalence, it remains hyperintensional in the weak sense that a posteriori equivalent sentences continue to have distinguishable epistemic n-contents. For instance, 'Hesperus is Hesperus' and 'Hesperus is Phosphorus' will always have distinguishable epistemic n-contents in our framework since $\mathcal{S}$ does not contain any rules that allow agents to infer facts about Hesperus from facts about Phosphorus, or vice versa.

Second, the relation of having indistinguishable epistemic n-contents becomes an identity relation once n approaches infinity. For in this limit, what is n-step inferable in $\mathcal{S}$ coincides with what is inferable in $\mathcal{S}$ *simpliciter*. As such, whenever A and B are a priori equivalent, the epistemic n-intension of $A \leftrightarrow B$ is necessary and hence true at all scenarios in W_n. It then follows that the epistemic n-intension of A is true at a given scenario w in W_n if and only if the epistemic n-intension of B is true at w. So the epistemic n-intensions of A and B will have the same truth-values at all scenarios in W_n when n approaches infinity.

Third, once n approaches infinity, epistemic n-intensions cut content in much the same way that Chalmersian primary intensions do.[11] On Chalmers' framework, the primary intension of a sentence represents what we might call its *ideal* epistemic content: roughly, the conditions under which the sentence is true at priori a closed and consistent scenarios. Two sentences A and B can then be said to have the same ideal epistemic content just in case $A \leftrightarrow B$ is "deeply epistemically necessary," where a sentence is deeply epistemically necessary just in case it is a priori for an idealized agent with no cognitive limitations (Chalmers 2011). So essentially, a sentence is deeply epistemically necessary just in case it is in principle inferable in $\mathcal{S}$. Thus epistemic n-intensions become extensionally equivalent with Chalmersian primary intensions when n approaches infinity. When n is finite, we can think of epistemic n-intensions as a kind of "non-ideal" primary intensions that represent what we might call the *non-ideal* epistemic content of sentences: the conditions under which sentences are true at the n-expansions of scenarios in W_n. There is thus a natural way in which we can think of epistemic n-intensions

[11] See Chalmers (2002; 2011) for details.

as non-ideal analogues of Chalmers' primary intensions.

4.1 Fregean cases revisited

Obviously, (Fregean thesis) does not give the right semantic verdicts in paradigmatic Fregean cases such as (1) to (6) when we focus on the limit cases where $n = 0$ or $n \to \infty$. Rather, we need an intermediate value of n for which sentence pairs such as (1) to (3) have distinguishable epistemic n-contents, whereas sentence pairs such as (4) to (6) have indistinguishable epistemic n-contents. While the appropriate value of n will depend on which specific inference rules we include in $\mathcal{S}$, nothing of importance will hinge on the choice here. Below we show how (Fregean thesis) can yield the correct semantic verdicts in (1) to (6) when $n = 2$ and $\mathcal{S}$ contains appropriate 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. $A \wedge A$.

Suppose that $\mathcal{S}$ contains the basic introduction and elimination rules from propositional logic. Given this, $A \wedge A$ is 1-step inferable from A by one application of conjunction introduction, and A is 1-step inferable from $A \wedge A$ by one application of conjunction elimination. So $A \leftrightarrow (A \wedge A)$ is 2-step inferable in $\mathcal{S}$, which, by Theorem 1, means that the epistemic 2-intension of $A \leftrightarrow (A \wedge A)$ is necessary.[12] So, by (Fregean thesis), A and $A \wedge A$ have indistinguishable epistemic 2-contents. By contrast, (3a) and (3b) have distinguishable epistemic 2-contents since (3b) is not 1-step inferable from (3a), and vice versa. So (Fregean thesis) can deliver the right semantic verdicts for (3) and (6).

[12]Depending on the exact specification of the logical fragment of $\mathcal{S}$, further steps might be required to infer $A \leftrightarrow (A \wedge A)$. For instance, if we adopt a Lemmon-style natural deduction system, $A \leftrightarrow (A \wedge A)$ would be 5-step inferable in $\mathcal{S}$ because we would need additional steps for making and discharging assumptions.

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. $y = x$.

We can reasonably stipulate that $\mathcal{S}$ contains a rule for commutativity of identity: $x = y \vdash_1 y = x$. Given this, $x = y$ is 1-step inferable from $y = x$, and vice versa. So $x = y \leftrightarrow y = x$ is 2-step inferable in $\mathcal{S}$, which, by Theorem 1 and (Fregean thesis), means that $x = y$ and $y = x$ have indistinguishable epistemic 2-contents. By contrast, (2a) and (2b) have distinguishable epistemic 2-contents since no reasonable specification of the mathematical fragment of $\mathcal{S}$ will make (2b) 1-step inferable from (2a), and vice versa. So again, (Fregean thesis) can deliver the right semantic verdicts for (2) and (5).

Consider, finally, the analytical 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.

We already know that (1a) and (1b) have distinguishable epistemic n-contents for any value of n. Moreover, we can ensure that (4a) and (4b) have indistinguishable epistemic 2-contents by including in $\mathcal{S}$ an inference rule of the following form: (the number of Fs is x) $\vdash_1$ (there are x many Fs), and vice versa. So again, (Fregean thesis) can deliver the right semantic verdicts for (1) and (4).

While we have illustrated, on a case-by-case basis, how our semantics can capture paradigmatic Fregean distinctions among contents, it remains an open question whether it is feasible to specify $\mathcal{S}$ in a way that adequately captures all sound a priori reasoning. 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. Also, the rules in $\mathcal{S}$ need not offer a *positive* characterization of the types of situations in which a given sentence can be inferred from a given set of premises. Rather, for some sentences, it may turn out that 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.

In general, the issues involved in specifying $\mathcal{S}$ are in many respects similar to issues that arise in discussions of the logical status of entailment patterns in natural languages. If we adopt a too restrictive view on what logical entailment is—for instance entailment in propositional logic—logical analysis 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 and more permissive notion of entailment. Questions concerning how to specify this broader notion of entailment, it seems to us, bear resemblance to questions concerning how to specify the (non-logical) rules in $\mathcal{S}$. If so, we might look to work in formal semantics to help us get clear on the kinds of rules that $\mathcal{S}$ should contain. 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.[13]

But whether or not it is ultimately reasonable to model a priori reasoning in terms of a step-based sequence of applications of a priori inference rules, let us repeat that the guiding idea behind our semantics does not rely on this assumption. In principle, our semantic framework is compatible with any quantitative measure of the complexity of different chains of a priori

[13]For detailed discussions of these matters, see Glanzberg (2015) and Zimmermann (1999).

reasoning. So those who want to represent complexity of reasoning in terms of time consumption, brain activity, or some other quantity, can still use a version of our framework.

4.2 Subsentential expressions

While our main focus in this paper concerns the individuation of sentential content, it is natural to wonder if epistemic n-intensions can also be used to individuate the contents of subsentential expressions such as singular terms, predicates, and the like. Consider the following pairs of subsentential expressions:

(7) a. '$4 + 1$'
b. '$1 + 4$'

(8) a. 'Bachelor'
b. 'Unmarried man'

(9) a. '7'
b. '$(5^2 \times 211 - 4) \div 753$'

(10) a. 'Hesperus'
b. 'Phosphorus'

On a Fregean picture, the pairs in (9) and (10) have distinguishable contents, whereas those in (7) and (8) do not. For instance, while it may be cognitively significant for an ordinary agent to learn that 'Hesperus' and 'Phosphorus' refer to the same object, the same cannot be said of '$1 + 4$' and '$4 + 1$'. This suggests that we individuate the contents of subsentential expressions in much the same way that we individuate the contents of sentences. More specifically, we propose the following generalization of (Fregean thesis):

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

Here the relation '≡' will vary in accordance with the type of expression in question: for singular terms the relation will be identity '=', for sentences it will be the material biconditional '↔', and so on.[14]

It is easily seen that (Generalized Fregean thesis) can yield the right semantic verdicts in cases such as (7) to (10). For instance, if $\mathcal{S}$ contains the commutative law for addition, the identity '$4+1=1+4$' is 1-step inferable in $\mathcal{S}$, which means that the epistemic 1-intension of '$4+1=1+4$' is necessary. By (Generalized Fregean thesis), then, '$4+1$' and '$1+4$' have indistinguishable epistemic 1-contents. By contrast, since it is not even in principle inferable in $\mathcal{S}$ that 'Hesperus' is identical to 'Phosphorus', (10a) and (10b) will have distinguishable epistemic n-contents, for any value of n.

While (Generalized Fregean thesis) gives us a way to individuate the semantic contents of subsentential expressions, we will not attempt to define the epistemic n-intensions of such expressions in detail. Generally, we want to understand such intensions as functions from scenarios in W_n to extensions, where the kind of extension in question will depend on the type of expression in question: the extensions of singular terms are objects, the extensions of predicates are sets of objects, and so on. However, to evaluate the epistemic n-intensions of subsentential expressions at scenarios in W_n, we need to know what it means to n-expand scenarios at the level of subsentential expressions. And since it is not immediately clear what it means to draw a priori inferences from subsentential expressions when those are not embedded in whole sentences, it is not immediately clear what it means to n-expand scenarios at the subsentential level. So while we expect that a plausible story can be told about how agents can reason a priori with subsentential expressions, we leave a detailed investigation for another occasion.

4.3 Jago on Fregean content

Mark Jago (2014) has recently proposed a hyperintensional semantics that—much like ours—promises to model Fregean content. Jago's hyperintensions are defined on worlds in an epistemic space, call it $W_{\mathcal{J}}$, which contains no

[14]We are here inspired by Chalmers (2006).

blatantly or trivially inconsistent worlds. Essentially, that is, $W_{\mathcal{J}}$ contains no worlds from which a contradiction can be easily or trivially inferred. According to Jago, it is a *vague* matter which inferences count as trivial, and hence a vague matter which worlds belong to $W_{\mathcal{J}}$. While a world that falsifies a trivial truth such as "$1 + 3 = 3 + 1$" is determinately not part of $W_{\mathcal{J}}$, it is indeterminate whether worlds that falsify more subtle inconsistencies are part of $W_{\mathcal{J}}$ (Jago 2014, ch. 7). But irrespective of how we draw the line between trivial and non-trivial inconsistencies, Jago allows that worlds in $W_{\mathcal{J}}$ may be arbitrarily partial or truth-gappy. So while no worlds in $W_{\mathcal{J}}$ falsify trivial logical truths such as $A \vee \neg A$, there are plenty of worlds in $W_{\mathcal{J}}$ that *fail* to verify such trivial truths.

Jago then holds that two sentences A and B "do not differ determinately in sense" just in case 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 Fregean content on Jago's account—and similarly for other trivially equivalent pairs of sentences. By contrast, if B is a complicated logical equivalent of A, A and B will determinately differ in Fregean content since there are worlds that verify A and falsify B that are determinately part of $W_{\mathcal{J}}$.

Why believe that semantic content individuation should be understood as a vague matter in the first place? Jago's motivation stems from an alleged vagueness at the level of knowledge and belief, which he captures by an epistemic logic designed to validate the following principle that we will call:

> **(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)

Jago proposes (Epistemic indeterminacy) as a solution to a problem—what he calls "the problem of rational knowledge" (Jago 2013b)—that arises if we want to model agents who fall short of logical omniscience, but who nevertheless are logically competent. The problem is that agents who are logically competent in the sense of knowing at least all the *trivial* logical consequences of their knowledge must also be logically omniscient. For it can be shown that agents who know all trivial logical consequences of what they know must thereby know all logical consequences of what they know, and hence be logically omniscient.[15] So, on pain of logical omniscience, Jago concludes that "it is possible to fail to know or believe trivial truths (and, more generally, trivial consequences of one's beliefs)" (Jago 2014, p. 243). But if agents can miss out on trivial logical consequences of their knowledge, it seems that we have merely traded logical omniscience for radical logical incompetence. To avoid this charge, Jago holds that it can never be *determinate* that an agent fails to know any particular trivial consequence of her knowledge. And since what is indeterminate is not rationally assertable, (Epistemic indeterminacy) is thus supposed to prevent us from treating logically non-omniscient agents as logically incompetent.

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. And since any two trivially equivalent sentences A and B do not determinately differ in Fregean content, on Jago's account, we get that A and B do not determinately differ in Fregean content just in case anyone who knows A do not determinately fail to know B, and vice versa. As far as we can tell, this biconditional claim captures Jago's motivation for treating Fregean content as a vague phenomenon. However, as we shall argue now, there are several problems with Jago's use of (Epistemic indeterminacy) in the epistemic context. And given that his motivation for treating semantic content individuation as vague relies on (Epistemic indeterminacy), these problems carry over to his account of Fregean content.[16]

First, (Epistemic indeterminacy) lacks independent motivation. Jago motivates the thesis by appeal to an alleged structural similarity between the problem of rational knowledge and the sorites paradox:

> 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

[15]See Bjerring (2013) and Jago (2013b; 2014) for detailed discussions of this result.

[16]In Bjerring & Rasmussen (ms.), we unfold the critique below in more detail. Here we rehearse the points that are most relevant to the discussion of Fregean content.

> 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. The tolerance conditionals for 'red', for example, say that (in a sorites series of colour samples), if sample n is red then so is sample $n + 1$. Clearly, not all such conditionals are true; but we cannot say or discover which is false. (Jago 2013b, p. 1155)

This alleged structural similarity between the problem of rational knowledge and the sorites paradox, we submit, *presupposes* rather than *supports* (Epistemic indeterminacy). For note that the similarity is brought out by the claim that "any attempt to say precisely at which point in the deduction the agent's knowledge gives out is doomed to failure" together with the claim that "any attempt to say precisely which tolerance conditional fails to hold is doomed to failure." But the former claim is plausible only if we already accept (Epistemic indeterminacy). For if we do not, nothing prevents us from pointing out precisely at which point in a deduction the agent's knowledge gives out. So there is only a structural similarity between the problem of rational knowledge and the sorites paradox if (Epistemic indeterminacy) is already assumed. As such, Jago has offered no independent reason to accept (Epistemic indeterminacy), and derivatively no independent reason to treat semantic content individuation as a vague matter.

Second, there are intuitive counterexamples to (Epistemic indeterminacy). Consider a simple, logically non-omniscient artificial agent who knows just the propositions $p_1, p_1 \to p_2, p_2 \to p_3, \ldots, p_{k-1} \to p_k$, and suppose the agent can apply modus ponens only once. Given this, there is exactly *one* trivial consequence of the agent's knowledge, namely p_2. In light of the problem of rational knowledge, as we saw above, we know that a logically non-omniscient agent must fail to know at least one trivial consequence of her knowledge. And since p_2 is the only trivial consequence of the agent's knowledge, it follows that the agent cannot know p_2. So we determinately know that the agent fails to know p_2. But this runs counter to (Epistemic indeterminacy).

One might object that the counterexample above is too simplistic since the agent only has very little knowledge and only a single inference rule available for deductive reasoning. Yet, we see no principled reason why an epistemic logic should not apply to such artifical agents. On the contrary, several philosophers have been interested in modelling epistemic agents with highly incomplete reasoning mechanisms and limited cognitive resources.[17] But even if we restrict our attention to complex agents, trouble continues to brew for (Epistemic indeterminacy). Consider a logically non-omniscient, yet logically competent agent such that there are m different trivial consequences $q_1, \ldots, q_m$ of the agent's knowledge. Suppose we now ask this agent whether q_1 is true, then whether q_2 is true, and so on all the way up to q_m. Given that the agent is logically competent, we can assume that she answers "yes" to each question. Yet, since the agent is logically non-omniscient, there must be at least one q_i that she fails to know. So we face an explanatory challenge: how can the agent affirmatively answer each question when she does not know all of q_1 to q_m? It is unclear how (Epistemic indeterminacy) can help us meet this explanatory challenge. The thesis merely says that the agent's lack of belief in some q_i must be indeterminate, but we are not offered an explanation of why the agent nevertheless can assent to each q_i despite not knowing all of q_1 to q_m.

So we conclude that (Epistemic indeterminacy) lacks both motivation and plausibility, and hence that Jago's hyperintensional account of Fregean content, which relies on this thesis, lacks equally in both respects.

[17]See, among others, Konolige (1986) and Fagin et al. (1995).

5 Applications

We will round off by illustrating how our semantic framework can be used to solve various philosophical problems that call out for a hyperintensional, yet non-trivial analysis. We will focus on two problems that have received considerable attention over the years: *the scandal of deduction* in the philosophy of information, and *the problem of logical omniscience* in formal epistemology. But we are optimistic that our framework can be applied across an even wider range of hyperintensional contexts.

5.1 The scandal of deduction

Worlds-based accounts of informational content treat a proposition as informative for an agent just in case it allows the agent to rule out any scenario that falsifies the proposition as a way the world might be—that is, as *epistemically impossible*.[18] Suppose I wonder whether it is snowing in Canberra. For all I know, we may or may not live in a world where it snows in Canberra. My reliable friend now tells me that it is indeed snowing in Canberra. This piece of testimony is informative because it allows me to rule out scenarios in which it is not snowing in Canberra as epistemically impossible. Similarly, suppose I wonder whether there are spatially extended objects without mass. For all I know, we may or may not live in a world where such objects exist. I now infer from experience that all spatially extended objects have a mass. This inductive inference is informative because it allows me to rule out scenarios in which some spatially extended objects are massless as epistemically impossible.

Can *deductive* reasoning be informative in this way? It surely seems so. While we currently expect that Goldbach's Conjecture is true, we do not know it: for all we know, we may or may not live in a world where the Conjecture is true. Suppose I manage to give a proof of the Conjecture from

[18]For ease of exposition, we will talk about propositions being informative *for an agent* rather than being informative *simpliciter*, but nothing of importance will hinge on this qualification. See D'Agostino & Floridi (2009) and Sequoiah-Graysonand (2008) for general discussion of accounts of information as exclusion of possibilities.

the Peano Axioms. This deductive inference is informative because it allows us to rule out scenarios in which there are even integers greater than 2 that are not expressable as the sum of two primes. Similarly, when Gödel proved the Incompleteness Theorems, this deductive reasoning was informative in part because it allowed us to rule out scenarios in which Hilbert's program was achievable. So given that we understand information in terms of elimination of epistemically possible scenarios, it seems clear that deductive reasoning can be informative.

Yet, standard intensional accounts of information render all deductive reasoning *uninformative*. For such accounts presuppose that all epistemically possible scenarios are *logically* possible: they verify the conclusion of any valid deductive inference whenever they verify its premises. As a result, there are no scenarios violating deductive inferences that agents can come to rule out as epistemically impossible. So all deductive reasoning is treated as uninformative in the intensional framework. Hintikka (1973) referred to this counterintuitive result as *the scandal of deduction*.

We can avoid the scandal by treating information as a *hyperintensional* phenomenon.[19] In particular, we can allow that some epistemically possible scenarios are logically *im*possible: they can verify the premises of a valid deductive inference but falsify its conclusion. By including logically impossible scenarios in our model, we have ensured that there are scenarios violating deductive inferences that agents can come to rule out as epistemically impossible. In turn, deductive reasoning can be treated as informative. For instance, we can appeal to the epistemic possibility of a scenario that verifies the Peano Axioms but not Goldbach's Conjecture to explain why the inference from the Axioms to the Conjecture can be informative.

Information, however, is not only a hyperintensional phenomenon, but also a *non-trivial* one: while the proof of Goldbach's Conjecture from the Peano Axioms strikes us as significant and informative, the same cannot be said about trivial inferences such as 'A therefore $A \wedge A$' or '$x = y$ therefore

[19] Jago (2013a) presents an account of informational content that bears resemblance to the account we present here. Yet, since Jago's account relies on his general theoretical framework, which we critized above, we will not pause to discuss his proposal here.

$y = x$.' This suggests that we should avoid an "anything goes" view on information: there should be substantial constraints on what counts as an informative deductive inference. In specifying *which* deductive inferences should count as informative, we can trade on our interpretation of Frege's equipollence principle and say that a valid deductive inference from A to B is informative just in case $A \rightarrow B$ is not n-step inferable in $\mathcal{S}$. This effectively ensures that only *non-trivial* or *unobvious* deductive inferences are treated as informative. We can spell out this proposal in terms of epistemic n-intensions as follows:

> **(Information principle)** A valid deductive inference from A to B is informative iff the epistemic n-intension of $A \rightarrow B$ is not necessary.

Since, by Theorem 1, the epistemic n-intension of $A \rightarrow B$ is necessary just in case $A \rightarrow B$ is n-step inferable in $\mathcal{S}$, (Information principle) can ensure that only sufficiently complex deductive inferences count as informative.

(Information principle) shares the epistemic aspect of Frege's equipollence principle: whether a deductive inference counts as informative depends on the cognitive resources that agents have available for deductive reasoning. By varying the value of n, (Information principle) gives rise to a whole spectrum of ways of drawing the line between informative and uninformative deductive inferences. When $n = 0$, all deductive inferences count as informative since nothing is 0-step inferable in $\mathcal{S}$. When n approaches infinity, no deductive inferences count as informative since all deductive inferences can in principle be carried out in $\mathcal{S}$. For intermediate values of n, some but not all deductive inferences count as informative. It is not hard to see, for instance, how an appropriate specification of the value of n will allow us to treat the inference from the Peano Axioms to Goldbach's Conjecture as informative but the inference from A to $A \wedge A$ as uninformative. So (Information principle) allows us to reason about information as a hyperintensional, yet non-trivial notion: it avoids the scandal of deduction without entailing, implausibly, that all deductive inferences—however trivial or obvious—are informative.

5.2 The problem of logical omniscience

According to the standard possible-worlds semantics for belief (or knowledge), which goes back to Hintikka (1962), an agent is said to believe a proposition A just in case A is true at all possible worlds that are doxastically possible for the agent. It is well-known that this semantics gives rise to *logical omniscience*: agents are described as believing all logical consequences of what they believe, including all logical truths. To see why, suppose that an agent believes A, and let B be any logical consequence of A. Given that all doxastically possible scenarios are *logically* possible, every doxastically possible scenario that verifies A will verify B. And since, by assumption, every doxastically possible scenario verifies A, every such scenario verifies B, in which case the agent believes B. So the possible-worlds model of belief implies that agents believe all logical consequences of what they believe.

Yet, ordinary agents such as you and me clearly fall short of logical omniscience. While I am confident that $2 + 2 = 4$, I do not have an honest clue about what 17^{43} precisely amounts to. And while I know that $A \vee \neg A$ is tautological, I am ignorant about whether $(\neg B \wedge C) \rightarrow \neg(A \leftrightarrow (B \vee \neg A))$ is tautological, contingent, or contradictory—at least prior to thinking hard about it. So if our aim is to model ordinary rational agents such as humans or intelligent machines, logical omniscience should be avoided. This is what Hintikka (1975) referred to as *the problem of logical omniscience*.

Avoiding logical omniscience, however, should not be our only aim: we should also avoid describing agents as entirely logically *ignorant* or *incompetent*. After all, while ordinary agents are logically non-omniscient, most are nevertheless capable of teasing out at least *some* logical consequences of what they believe. If I believe that it is raining and that rain makes the streets wet, I will manage—at least in normal circumstances—to conclude that the streets are wet. So although the belief states of ordinary agents are not *closed* under logical consequence, such states nevertheless display substantial logical structure.

This suggests that belief—much like Fregean content and information—should be thought of as a hyperintensional, yet non-trivial phenomenon: it is hyperintensional because ordinary agents do not believe *all* logical consequences of what they believe, and non-trivial because such agents can at least come to believe *some* such consequences. In specifying *which* logical consequences we can expect ordinary agents to come to believe, we can once again trade on our interpretation of Frege's equipollence principle and say that whenever A is n-step inferable in $\mathcal{S}$ from what an agent believes, the agent can come to believe that A. This effectively ensures that agents can come to believe at least all the *trivial* or *obvious* a priori consequences of what they believe.

While it would take us too far afield to develop all the formal details of an epistemic logic that captures the proposal above, we want to give at least a rough sketch of how one might use principles such as (n-consistency) and (n-expansion) to erect a hyperintensional, yet non-trivial semantics for belief.[20] Initially, we replace the standard belief operator by an operator Bel^n with the following intended interpretation:

> $Bel^n(A)$: the agent believes A after some n steps of a priori reasoning.

When $n = 0$, we can think of Bel^n as a standard belief operator: what you believe after 0 steps of a priori reasoning just is what you believe (assuming, as we do, that nothing is 0-step inferable in $\mathcal{S}$).

For the semantics of Bel^n, let R be an accessibility relation consisting of a set of pairs of scenarios in W_n, and let V be an evaluation function from sentences and scenarios to truth-values. We can then give Bel^n the following truth-conditions:

> **($\boldsymbol{Bel^n}$ semantics)** $V(Bel^n(A), w) = T$ if and only if $V(A, v^n) = T$ for all $v : \langle w, v\rangle \in R$.

Informally, this semantics says that an agent believes A after some n steps of a priori reasoning just in case A is true at the n-expansion of each doxastically possible scenario for the agent. When $n = 0$, (Bel^n semantics) reduces to an impossible worlds semantics for a standard belief operator: $Bel^0(A)$ is true at w just in case A is true at all scenarios that are accessible from w.

[20] For a detailed exposition, see Bjerring & Rasmussen (ms.).

The following result shows that agents can come to believe every sentence that follows within n steps of a priori reasoning from what they already believe:

Theorem 2. If $Bel^0(A)$ and $A \vdash^n_S C$, then $Bel^n(C)$.

Proof. Let A be any proposition, let w be any scenario in W_n such that $V(Bel^0(A), w) = T$, and let C be any proposition such that $A \vdash^n_S C$. By (Bel^n semantics), $V(A, v^0) = T$ for all $v : \langle w, v \rangle \in R$. By ($n$-expansion), $V(C, v^n) = T$ for all $v : \langle w, v \rangle \in R$. So, by ($Bel^n$ semantics), $V(Bel^n(C), w) = T$.

By varying the value of n, (Bel^n semantics) allows us to reason about a whole spectrum of agents with different capacities for a priori reasoning. When $n = 0$, (Bel^n semantics) models agents who are completely logically incompetent: they cannot infer any a priori consequences of what they believe. That is, for any A and C such that C is an a priori consequence of A, the semantics allows that $Bel^0(A)$ is true while $Bel^0(C)$ is false. When n approaches infinity, (Bel^n semantics) models agents who are logically omniscient: they can infer all a priori consequences of what they believe. For intermediate values of n, (Bel^n semantics) models agents who are logically non-omniscient, yet logically competent: they can come to believe some, but not all logical consequences of what they believe. Thus (Bel^n semantics) gives us a hyperintensional, yet non-trivial account of belief: it steers clear of logical omniscience without entailing, implausibly, that agents are entirely incapable of engaging in logical reasoning.

6 Summary

We have presented a theory of semantic content that validates Frege's equipollence principle, and we have shown how the semantics successfully accommodates content that—much like Fregean senses—is both hyperintensional and non-trivial. The semantics is hyperintensional because it associates distinguishable contents with some necessarily equivalent sentences, and non-trivial because it associates indistinguishable contents with some such sentences. More specifically, on our view, two sentences A and B have indistinguishable epistemic n-contents just in case the epistemic n-intension of $A \leftrightarrow B$ is necessary, where the epistemic n-intension of $A \leftrightarrow B$ is necessary just in case $A \leftrightarrow B$ is true at the n-expansion of all n-consistent scenarios. By varying the value of n, our semantics provides us with a whole spectrum of notions of content with different levels of granularity.

While we have shown that epistemic n-content can do important semantic work, we acknowledge that it might not be useful for all semantic purposes. For example, since the individuation of semantic content in our framework varies with cognitive resources, epistemic n-content is not obviously useful for reasoning about communication or entailment patterns in natural language. To illustrate, suppose Brit has significantly more cognitive resources available for a priori reasoning than Anthon, and imagine that Brit says 'the interior angles of a hexagon in Euclidian space sums up to 4π' to Anthon. Because of the differences in cognitive resources, Brit and Anthon will associate different epistemic n-contents with this sentence. But if communication requires the transmission of a unique semantic content, we seem forced to say, implausibly, that there is no communication going on between Brit and Anthon. So if we want to account for how language is used in communication, epistemic n-content is not immediately useful. But this does not show that epistemic n-content is not a useful notion of semantic content altogether. While the task of accounting for communication is surely important, it is hardly the only important task in semantics. Semantics is also centrally involved with accounting for how language conveys information and expresses thought contents, and, as we have argued, epistemic n-content can do important work in these respects.

References

Bjerring, Jens Christian (2010). "Non-Ideal Epistemic Spaces". PhD thesis. Australian National University.

Bjerring, Jens Christian (2013). “Impossible Worlds and Logical Omniscience: An Impossibility Result”. In: *Synthese* 190.13, pp. 2505–2524.

— (2014). “Problems in Epistemic Space”. In: *Journal of Philosophical Logic* 1, pp. 1–18.

Bjerring, Jens Christian and Mattias Skipper Rasmussen (ms.). *A dynamic solution to the problem of logical omniscience*.

Bjerring, Jens Christian and Wolfgang Schwarz (forthcoming). “Granularity Problems”. In: *Philosophical Quarterly*.

Burge, Tyler (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.

Chalmers, David J. (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.

D’Agostino, Marcello and Luciano Floridi (2009). “The Enduring Scandal of Deduction”. In: *Synthese* 167.2, pp. 271–315.

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.

Fagin, R. et al. (1995). *Reasoning About Knowledge*. MIT Press.

Frege, Gottlob (1984). *Collected Papers on Mathematics, Logic, and Philosophy*. Oxford: Basil Blackwell.

Glanzberg, Michael (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, Jaakko (1962). *Knowledge and Belief*. Ithaca, N.Y.,Cornell University Press.

Hintikka, Jaakko (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, Mark (2013a). “The Content of Deduction”. In: *Journal of Philosophical Logic* 42.2, pp. 317–334.

— (2013b). “The Problem of Rational Knowledge”. In: *Erkenntnis* 6, pp. 1–18.

— (2014). *The Impossible: An Essay on Hyperintensionality*. Oxford University Press.

King, Jeffrey C. (2016). “Structured Propositions”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2016.

Konolige, K. (1986). *A Deduction Model of Belief*. Morgan Kaufman.

Kripke, Saul A. (1980). *Naming and Necessity*. Harvard University Press.

Laan, David A. Vander (1997). “The Ontology of Impossible Worlds”. In: *Notre Dame Journal of Formal Logic* 38.4, pp. 597–620.

Lewis, David (2004). “Letters to Priest and Beall”. In: *The Law of Non-Contradiction*. Oxford University Press, pp. 176–177.

Nolan, Daniel (1997). “Impossible Worlds: A Modest Approach”. In: *Notre Dame Journal of Formal Logic* 38.4, pp. 535–572.

Penco, Carlo (2003). “Frege: Two Theses, Two Senses”. In: *History and Philosophy of Logic* 24.2, pp. 87–109.

Priest, Graham (2005). *Towards Non-Being*. Clarendon Press.

Putnam, Hilary (1975). “The Meaning of ’Meaning’”. In: *Minnesota Studies in the Philosophy of Science* 7, pp. 131–193.

Schellenberg, Susanna (2012). “Sameness of Fregean Sense”. In: *Synthese* 189.1, pp. 163–175.

Sequoiah-Grayson, Sebastian (2008). “The Scandal of Deduction”. In: *Journal of Philosophical Logic* 37.1, pp. 67–94.

Stalnaker, Robert (1984). *Inquiry*. Cambridge University Press.

Stanley, Jason (2010). “"Assertion" and Intentionality”. In: *Philosophical Studies* 151.1, pp. 87–113.

Zalta, Edward N. (1997). "A Classically-Based Theory of Impossible Worlds". In: *Notre Dame Journal of Formal Logic* 38.4, pp. 640–660.

Zimmermann, Thomas Ede (1999). "Meaning Postulates and the Model-Theoretic Approach to Natural Language Semantics". In: *Linguistics and Philosophy* 22.5, pp. 529–561.