HYPERINTENSIONALITY

MSc Thesis *(Afstudeerscriptie)*

written by

Iliana Gioulatou
(born June 23rd, 1992 in Athens, Greece)

under the supervision of **Prof. Franz Berto** and **Dr. Luca Incurvati**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam*.

Date of the public defense:	**Members of the Thesis Committee:**
August 30, 2016	Dr. Maria Aloni
	Prof. Benedikt Löwe
	Dr. Soroush Rafiee Rad
	Prof. Frank Veltman

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

In the current thesis I examine the literature in hopes of arriving at a good, hyperintensional theory of content. A good theory of content, it is argued, is one that can draw certain hyperintensional distinctions, but of course, achieving a certain finess of grain does not, on its own, a good theory make. With these criteria in mind, I critically engage with the literature, and in particular with structured propositions accounts, extended possible world semantics, truthmaker semantics *à la* Stephen Yablo, and truthmaker semantics *à la* Kit Fine. I notice that there are two axes along which hyperintensionality runs – that of the impossible and that of aboutness – and the better theories are found, on the one hand, a possible-worlds based account born out of the works of Jago and Yablo combined, and Fine's state-based truthmaker semantics on the other, the question of who takes the trophy home being a matter of perspective.

Acknowledgements

There's more people I should thank than I have space to do so. First and foremost, I want to thank my supervisors Franz and Luca; Franz for, other things aside, giving me incentive to, allowing me to follow, and helping me develop strong philosophical interests that lead to this thesis, and of course for getting Luca on board, without whose help, advice, comments and direction I would be submitting a considerably inferior piece of work. I wish to thank many friends for their support: Bas, Stephen, Albert and Tom for helpful and illuminating discussions, Dina, Mira, Bertrand, and Charlie for, among other things, putting up with many a rants, Thomas for lending a hand (holding keys to an apartment) at a time of crisis, and Antigone for distracting me as she always does. I owe gratitude to the Onassis Foundation for their generous and vital support throughout my second year of studies, and to the E.W.Beth Foundation for doing so during my first. Lastly, mum, this is the last line of the thesis you have to read (although I know you will keep going): thank you for everything.

Contents

1 Introduction **4**
1.1 Motivation . . . 4
1.2 Outline . . . 5
1.3 The Data . . . 6

2 Structured Propositions **9**
2.1 Introduction – Outline . . . 9
2.2 Russellian Propositions . . . 10
2.2.1 Soames on the Main Tenets . . . 10
2.2.2 King on Structure . . . 14
2.3 Frege's Puzzle and Belief Reports . . . 18
2.3.1 Russellianism . . . 20
2.3.2 Crimmins and Perry . . . 26
2.3.3 Chalmers . . . 27
2.4 Conclusion . . . 31
2.4.1 The Data and the Verdict . . . 31

3 Extended Possible World Semantics **35**
3.1 Introduction – Outline . . . 35
3.2 Priest . . . 36
3.2.1 Semantics for $\mathcal{L}$ . . . 36
3.2.2 Semantics for $\mathcal{L}^+$ . . . 44
3.3 Jago . . . 52
3.3.1 The Problem of Bounded Rationality . . . 52
3.3.2 The Epistemic Space . . . 54
3.3.3 Modelling Content . . . 56
3.4 Conclusion . . . 60
3.4.1 The Data and the Verdict . . . 60
3.4.2 The Metaphysics of Impossible Worlds . . . 63

4 Yablo **66**
4.1 Introduction – Outline . . . 66
4.2 Thick Propositions . . . 66
4.2.1 What they are . . . 66
4.2.2 Recursive Truthmaking . . . 70
4.3 Conclusion . . . 74
4.3.1 The Data . . . 74
4.3.2 The Verdict . . . 76

4.4 Yablo and the Impossible 77

5 Fine **81**
5.1 Introduction – Outline 81
5.2 Fine's Truthmaker Semantics 82
5.2.1 Semantics on a State Space 82
5.2.2 Semantics on an Extended State Space 84
5.3 Conclusion 87
5.3.1 The Data 87
5.3.2 The Verdict 89

6 Conclusion **92**

Chapter 1

Introduction

1.1 Motivation

An operator **R** is *intensional* just in case **R***A* and **R***B* are truth-conditionally equivalent whenever *A* and *B* are logically equivalent, but need not be when *A* and *B* are merely truth-conditionally equivalent. It is in this sense that intensional operators can distinguish between sentences that have the same truth value. However, such operators cannot distinguish between logically equivalent sentences. Operators that *can* are called *hyperintensional*. Within standard possible worlds semantics, hyperintensional operators cannot be defined, and hyperintensional notions cannot be captured.

Standard possible world approaches to modelling epistemic notions such as knowledge, belief, and informativeness often face difficulties that can be traced back to the failure of the underlying framework to distinguish between necessary equivalents. Logical omniscience is a paradigmatic case in point: an agent knows *A* just in case *A* is true in all epistemically accessible possible worlds. It thus follows that, given logical truths are true in all possible worlds, all agents know all logical truths, and assuming – as many do – that mathematical truths are also necessary, all agents know all mathematical truths as well. What is more, an agents knows all logical consequences entailed by what she knows, so given I know the rules of chess I would seem to know all winning strategies too. To move on to belief, given that agents often hold contradictory beliefs, and assuming classical logic, an agent holding a contradictory set of beliefs actually believes everything. Similarly strikingly so, since informativeness is captured by the difference between an agent's epistemically accessible set of worlds and the set of worlds that remain epistemically accessible once a new piece of information is added to the set of beliefs of the agent, logical and mathematical truths, as well as metaphysical necessities, fail to be informative; there is no epistemic gain whatsoever in devoting ones life to the study of logic, or mathematics.

Similar issues arise if we turn to *hope* or other propositional attitudes, but there are issues that arise even if we do not concern ourselves with intentional or epistemic notions. Under standard possible world semantics, a proposition is identified with a set of possible worlds, the worlds in which the proposition is true.[1] But a proposition is meant to capture the content of a sentence or utterance of a sentence expressing it, and many aspects of what we would naturally consider part of content are lost under the possible worlds picture. For example, all logical and mathematical truths have the exact same content, as do all impossibilities. But intuitively,

[1]See for example Stalnaker 1976.

the content of '2+2=4' is different from that of '$A \vee \neg A$', as is the content of '$A \wedge \neg A$' different from that of '2+2=5', if only because the sentences of each pair do not seem to concern the same things.

It seems like intentional and epistemic notions, as well as semantic content cry out for hyperintensional analyses. But there is further reason to want to go hyperintensional. Nolan reminds us that the last few decades of the twentieth century witnessed an intensional revolution as "[m]any metaphysicians rejected the doctrine, associated with Quine and Davidson, that extensional analyses and theoretical resources were the only acceptable ones (Nolan 2013, p.1)", resulting in one of the most fruitful eras of late modern analytic philosophy. And we are currently witnessing a hyperintensional revolution in its infancy, so to speak: more and more philosophers are coming up with hyperintensional analyses of essence (e.g.Fine 1994), explanation (e.g.Schnieder 2011), metaphysical grounding (e.g. Fine 2012, Schaffer 2009), what it means for a property to be *intrinsic* (e.g. Eddon 2011), disposition ascriptions (e.g. Jenkins and Nolan 2012) and many more. And although hyperintensional metaphysics is still in its infancy, there is justified hope that many a traditional metaphysical issues would greatly benefit from a hyperintensional analysis.

1.2 Outline

In the following introductory section I expand on the limitations of standard possible world analyses, so that I conclude on a list of data that I consequently use to demarcate desirable hyperintensional distinctions. Given that standard possible world semantics (SPW) is faced with such limitations, there seems to be three ways to go: (1) deny there is a real problem, (2) extend or reinterpret the possible worlds framework so as to cut more finely, or (3) abandon possible worlds altogether and opt for a more fine-grained framework to begin with. I dismiss denial – as for example manifested in Stalnaker's work on belief (see for example Stalnaker 1987) – and take it that the discussion of the following section supports the claim that content is hyperintensional, not to mention that, as argued, welcoming hyperintensional analyses would benefit the field across the border.

In the main body of the thesis comprising of chapters 2 to 5, I examine, compare, contrast and evaluate against the data four different theories of content that have made a claim to being sufficiently hyperintensional: structured accounts of propositions – falling under (3) – in Chapter 2, extended possible world semantics in Chapter 3 and a theory of meaning based on Yablos work in Yablo 2014a in Chapter 4– both falling under (2) – and lastly, in chapter 5, Fine's work on truthmaker semantics – falling under (2). My aim is to discover which, if any, of the accounts is more satisfactory, based on two criteria: the first one is that the account arrives at contents that are individuated finely enough – to which end I use the data – and the second and broader one is that the account is theoretically virtuous – the idea behind which is that an account of meaning that suffers a serious theoretical setback – that is internally inconsistent, or severely underdeveloped for example – is not a good account of meaning, let alone a good hyperintensional account of meaning. Each of the main body chapters is thus loosely structured as follows: I introduce the account under consideration, examining its main tenets, and noting any issues that might be taken with those. I then present semantic clauses for a first order language and examine the relevant rendering of the pairs of data, taking note where the account conflates propositions. I conclude by summarising how well the account fairs with the data, and whether it faces any problems of a theoretical nature. As soon as I have my four conclusions, one for each chapter of the main body, I should be able to conclude which, if any, of the theories is a good, hyperintensional theory of content.

Of course, things turn out to be not as simple as that, for as soon as Chapter 3 we see a certain pattern: hyperintensionality seems to develop across two axes, one of the impossible and one of aboutness, and a theory achieves the best score on the data-test when it manages to incorporate both. Thus in the end two theories, an account I arrive at by combining the works Yablo (who works towards aboutness) and Jago (who works towards the impossible), and Kit Fine's truthmaker semantics have virtues and vices that cannot be weighed as straightforwardly as one would hope.

1.3 The Data

According to the received view, propositions are sets of possible worlds, or characteristic functions thereof. Once this thesis is coupled with standard possible world analyses of epistemic notions, as well as purely *relational* analyses of propositional attitudes in general, the following are hard to account for:

Necessary Equivalents

1. Necessary truths – Mathematical or logical:

$$2+2=4 \tag{1.1}$$
$$3+3=6 \tag{1.2}$$
$$e^{i\pi}=-1 \tag{1.3}$$
$$P \vee \neg P \tag{1.4}$$

   Intuitively, (1.1), (1.2), (1.3) and (1.4) express different propositions, an intuition one could justify by appealing to the fact that they all seem to be about different things. Nevertheless, according to the current view – assuming classical logic and that mathematical truths are necessary – they express one and the same proposition.

2. Other necessary truths – Frege:

$$\text{All } A\text{'s are } B\text{'s.} \tag{1.5}$$
$$\text{All } A\text{'s are } A\text{'s.} \tag{1.6}$$
$$\text{H is P.} \tag{1.7}$$
$$\text{H is H.} \tag{1.8}$$

   On the assumption that all A' are B's and that H is P, the theory yields that (1.5), (1.6) express the same proposition, and that (1.7), (1.8) express the same proposition. However, as per Frege's Puzzle, there are strong reasons to believe that this is not so: for example, (1.7) might be informative while (1.8) not so.[2]

3. Necessary falsehoods:

$$\text{Amy squared the circle.} \tag{1.9}$$
$$\text{Bob squared the circle.} \tag{1.10}$$
$$2+2=5 \tag{1.11}$$
$$P \wedge \neg P \tag{1.12}$$

   Assume classical logic and that mathematical truths are necessary. Then, according to SPW semantics, all of the above express one and the same proposition – the null set. But this result, as was the case with necessary truths, is highly counterintuitive.

[2]For a variant of Frege's puzzle generated by (1.5) and (1.6), see Ripley 2012.

Propositional Attitude Reports

1. Knowledge and necessary truths

$$\text{Amy knows that (1.1).} \tag{1.13}$$

$$\text{Amy knows that (1.3).} \tag{1.14}$$

The underlying problem is that of logical omniscience: according to the SPW rendering of knowledge, and the SPW account of propositions, all agents know all necessary truths. But this seems false: someone who dropped math in high-school does not know as many theorems as a mathematician does. Moreover, SPW analyses of knowledge and belief have it that all agents know all logical consequences of what they know. For example, if we know the first order Peano axioms then we apparently know whether Goldbach's conjecture is true. This is absurd. More generally, identifying propositions with sentential intensions predicts a certain sort of logical closure which we simply do not have. But the problem resurfaces on the level of propositional attitude reports, too, where the substitution of a that-clause for a co-intensional is not always perceived to be *salva veritate*, contra the SPW prediction. Take Amy for example, a smart 5 years old who has already learned to add small numbers, but has no idea about e or π or negative numbers for that matter. It seems then that (1.13) is true while (1.14) is false, but the current theory predicts they must have the same truth value – true.

2. Frege's Puzzle and Belief reports:

$$\text{A believes that (1.7).} \tag{1.15}$$

$$\text{A believes that (1.8).} \tag{1.16}$$

One can build a case such that strong intuitions dictate that (2.7) is false while (2.8) is true. However, given that H is P and that names are rigid designators, the current theory yields that the sentences are truth-conditionally equivalent.

3. Necessary truths and other propositional attitude reports:

$$\text{A is surprised that (1.5).} \tag{1.17}$$

$$\text{A is surprised that (1.6).} \tag{1.18}$$

According to the possible worlds account, (1.17), (1.18) must have the same truth value since (1.5) and (1.6) are the same proposition. Yet – as Ripley does in (Ripley 2012) – as strong a case as for Frege's Puzzle can be constructed such that intuitions dictate that the former is true while the latter false.

4. Belief and Necessary Falsehoods: If an agent believes a necessary falsehood, then the theory predicts that for any proposition P the agent believes P. Again, this is if possible world semantics are paired with classical logic. However, they usually are and this is yet another unpalatable result.

Conditionals Many theorists – of the PW but of other persuasions, too – have often supplemented their accounts, taking propositional attitudes to be special cases and thus responding some way or another to variants of Frege's Puzzle. However, there seem to be issues that persist and can be brought out by embedding intuitively distinct co-intensional propositions in contexts that are traditionally considered to not be opaque; if we are dealing with conditionals for example, such supplements fail to be of any help by design. And Ripley (Ripley 2012) argues that the problem arises whenever clauses are embedded.

1. Indicative conditionals:

$$\text{If H is not P, then H is not P.} \tag{1.19}$$
$$\text{If H is not P, then H is not H.} \tag{1.20}$$

   Assuming that H is P, then both the antecedent and the consequent of both conditionals above are one and the same, and hence both conditionals are predicted to have the same truth value. However, it has been argued by Ripley (Ripley 2012) and Jago (Jago 2014) that, intuitively, the former is true while the latter is false.

2. Subjunctive conditionals:

$$\text{If Amy had squared the circle, then Amy would be famous.} \tag{1.21}$$
$$\text{If Bob had squared the circle, then Amy would be famous.} \tag{1.22}$$

   According to Ripley and Jago, intuitively (1.21) is true while (1.22) is false, but it is hard to see how a theory that treats the antecedents of both conditionals as one and the same proposition can accommodate for this.

3. Counterpossible conditionals. Even adopting the Lewis-Stalnaker's analysis of counterfactuals, counterpossible conditionals prove problematic since they all come out trivially true.[3] This is not a good prediction for we often reason from impossible antecedents, as for example when discussing competing mathematical or logical theories.

Pairs of the above sentences will serve as my data in the chapters to come, and be sure that when brought about a more lengthy discussion as to whether and as to why they ought to express distinct propositions will be supplied. I wish the current section to serve as an introduction, as a rough guide, and as a place the reader can return to if they need to recollect the sum of the data that are coming the theories' ways.

[3] See for example Lewis 1973, Stalnaker 1968.

Chapter 2

Structured Propositions

2.1 Introduction – Outline

Sentences express propositions – or so the story goes.[1] It is by virtue of expressing the same proposition, for example, that distinct sentences – perhaps sentences of different languages – can succeed in 'saying the same thing'. Among proponents of propositions, it is for the most part uncontroversial what functions these entities are meant to perform; theorists agree that propositions are the bearers of truth-values and modal properties, as well as the objects that agents bear cognitive attitudes towards.[23] It is propositions that can be true or false, contingent, possible or necessary, and believed, known or doubted by epistemic agents.

Propositions are often characterised by these roles they are meant to perform, but this leaves an important question unaddressed: what is the metaphysical nature of propositions? And there is indeed little consensus as to what sort of entity a proposition is. However, as we shall see, whether the nature a given account attributes to propositions is such that the entities in question can fulfill the roles they are all along envisioned to play, has considerably shaped the debate on the latter metaphysical issue. In other words, the question we seem to be attempting an answer to is this: assuming there is a single sort of entities that function as bearers of truth-values, possessors of modal properties and subjects of propositional attitudes, what is their metaphysical status?

A very prominent family of views identifies propositions with sentence intensions, that is, with unstructured set-theoretic constructions consisting of truth-supporting circumstances. More precisely, according to the simple possible worlds account, a proposition P *is* the set of possible worlds in which P is true, or equivalently, it *is* the function from the set of all possible worlds to truth values that maps a world w to *True* if and only if P is true in w. However, as we have seen, there are various problems that SPW faces which ultimately trace back to the failure of the theory to deliver entities that are fine-grained enough to fulfill the agreed upon roles

[1] For now I leave it aside whether sentences express propositions in a given context, or whether speakers express propositions by uttering sentences in context.

[2] Of course there are dissidents. Lewis 1980 for example, doubts whether propositions are the right entities to serve as contents of sentences, but there is general consensus enough to black-box such matters. As Bealer puts it, "[t]his paper begins in the middle of a long story. To tell my part of the story, I will need to assume the central tenets of the traditional theory of propositions (Bealer 1998, p.1)" For the tenets see both Bealer 1998 and for a longer discussion of those §2 and 3 of McGrath 2014

[3] There are of course people who deny the existence of propositions as for example Jubien 2001, and others who express more moderate skepticism such as Hofweber 2006, but I will be taking throughout a realist stance towards propositions.

propositions are meant to play. The coarseness of the intensional approach was readily noticed, and with Kripke's seminal *Naming and Necessity* (Kripke 1980) bringing back to prominence an essentially Millian account of singular terms, direct reference theory was coupled with the metaphysical claim that propositions are structured entities whose constituents and structure mirror that of the sentences expressing them. This resulted in a broadly Russellian picture of propositions that aspired to deliver the desired finess of grain.

The Russellian approach has been defended by theorists such as Soames, Salmon and King and it is the structuralist approach I mainly focus on. In what follows, I present the basic tenets of the view as put forth by Soames (Soames 1985, Soames 1987) (§2.2.1), follow King (§King 1995) in explaining what the *structure* of a structured proposition amounts to (§2.2.2), and look at Salmon's (§Salmon 1986) and Soames's (Soames 2002) defenses of some *prima facie* counter-intuitive consequences of the account (§2.3). The views of these theorists do indeed diverge at points but it is the conglomeration I describe above that I will be referring to as 'the' Russellian approach; it cannot be rightfully attributed to any one theorist in particular but I have attempted to navigate the large literature and "pick-and-mix" so as to put forward what I consider to be the best foot of Russellianism. As soon as certain issues with the Russellian approach are diagnosed, I more briefly look into alternative structuralist accounts such as the one put forth by Crimmins and the Fregean account defended by Chalmers (Chalmers 2011), in §2.3.2 and §2.3.3 respectively. In the final section, I re-examine how much fine-grained content each account is successful at delivering and look at some objections raised against structuralism, in the hope of ultimately determining whether among these views we have an overall well-motivated and satisfactorily hyperintensional theory of content (§2.4).

2.2 Russellian Propositions

2.2.1 Soames on the Main Tenets

In his 1985, and 1987 papers, Scott Soames launches a powerful attack against accounts that identify propositions with sets of truth-supporting circumstances. His aim is not limited to coarse-grained circumstantialism as exemplified by the SPW; to the contrary, his criticism is meant to extend to theories that identify propositions with more fine-grained entities such as sets of truth-supporting circumstances that may be inconsistent, incomplete or metaphysically impossible.[4]

Letting Soames summarise his own argument:

> If direct reference is possible and propositional attitude verbs have a relational semantics ... then the semantic values of sentences (objects of propositional attitudes) cannot be collections of truth-supporting circumstances.(Soames 1985, p.63).

What Soames essentially does is argue that, other than the usual troubles with SPW – some of which can be overcome by turning to more fine-grained circumstantialism – as soon as a circumstantialist admits that at least some singular terms – be they names, demonstratives, or variables – are directly referential, then together with a few other highly plausible assumptions, her view entails a whole new class of problematic cases that she cannot shake off by turning to inconsistent, incomplete or metaphysically impossible truth-supporting circumstances. And,

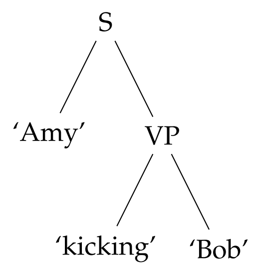
[4] This argument will be discussed extensively in the chapter about impossible world semantics. Soames is originally arguing against situation semantics – in particular the framework developed in Barwise's and Perry's *Situations and Attitudes* – but his argument will be shown unsuccessful against Priest's account. More details to follow, but for now, a short discussion of the criticism serves the narrative and makes explicit the motivation for Russellian propositions.

to be sure, Soames is a direct reference theorist and thus takes the argument to be a *reductio* of circumstantialism.[5] Soames concludes that what is needed is a completely new conception of semantic content disparate from sets of truth-supporting circumstances.

The positive account Soames puts forth is of course not “new" in the sense of having no historical precedent; as the name suggests, the theory he builds towards is a very close relative of the structured propositions the Russell of *Principia* endorsed. According to Soames – and most other structuralists for that matter – propositions are structured entities, made up of constituents bound together in a certain way. This is as of yet very uninformative for it tells us not what the constituents of propositions are, what determines the structure in which they figure, and what metaphysical ‘glue’ binds the parts together. All but the latter question which is postponed until §2.2.2 are addressed shortly. But first an auxiliary issue that brings us closer to the particulars of the view.

To say that propositions are structured is to state a purely metaphysical thesis. But as has been argued, there are certain functions that most want propositions to be able to perform, such as that they serve as the things that sentences express in contexts, or that speakers express in contexts with sentences. So, given that sentences have a syntactic structure and syntactic constituents, the question naturally arises of whether a proposition exhibits a structures and mereology that is somehow determined by the sentence expressing it. Rusellians, and all structuralist we shall be concerned with alike, argue that indeed the structure of a proposition can almost immediately be inferred from the LF structure of the sentence expressing it, and that the constituents of a proposition are the semantic values of its LF constituents.[6]

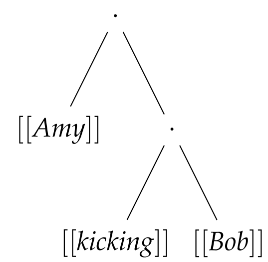
To better understand the claim, let us take an example. The simple sentence ‘Amy kicks Bob’ has the following syntactic tree:



What the structuralists now claim – all of the theorists we shall be considering but possibly not those who take an algebraic structuralist approach (see for example Bealer 1998) – is that we can get the structured proposition corresponding to this sentence by simply substituting the lexical items in the nodes with their semantic values. That is, representing the semantic value of an arbitrary lexical item l by $[[l]]$, we have the following:

[5]Salmon is a direct reference theorist too, while King neither explicitly endorses nor condones the dogma. However, when speaking of the Russellian account, I will be assuming that the direct reference thesis is part of the theory.

[6]Structuralists appeal to LF and not to surface grammatical structure for the usual reasons, to dispense with scope ambiguities and the like. For a more detailed discussion see King 1995.



We can represent this structure equivalently using nested sequences as in

$$\langle [[Amy]], \langle [[kicking]], [[Bob]] \rangle \rangle,$$

and indeed, I will be using the equivalent in-line representation whenever I can from now on, if only to save real trees.

Two main points remain: firstly, the structuralist needs to tell us what sort of semantic value she attributes to lexical items from the basic categories, and the rest of the current section is devoted to the Russellian response. Indeed the two alternative accounts we will be examining in §2.3.2 and §2.3.3 start diverging from Russellianism at this very point. And secondly, it seems that the structuralist is not quite done with structure. Substituting the lexical items at the nodes of the trees by their semantic values is of great help to get intuitions going, but metaphysically speaking, the account is incomplete; nothing has been said about the nature of the relation 'housing' propositional structure by binding together propositional constituents; the problem(s) of the unity if propositions we shall be concerned with in the next section.

Russellians – Soames and Salmon explicitly – are direct reference theorists: names, demonstratives, singular pronouns and variables are under their lights all devices of direct reference. The content of a directly referential term n with respect to some context c, is, simply, the object o n refers to in c. This is the gist of Kaplan's theory of indexicals (see Kaplan 1978), and direct reference theorists extend it to proper names, stripping – in a post-Kripkean spirit – the latter of any descriptive content they were once embellished with. As a result, according to this view names are rigid designators – they designate the same object in every world in which the object exists.[7] What is more, according to Russellians, the semantic value of an n-ary predicate just is the property or relation the predicate expresses. To take an example, the semantic value of "red" is not its possible world intention; it is not the function from worlds to sets of red individuals, it is simply the property of redness. Roughly speaking these are the main Russellian claims with respect to semantic values of constituents, but to get all the details straight and extend to the semantic contents of sentences we follow Soames in considering a simple first order language.

Consider a simple first order language $\mathcal{L}$ with logical symbols $\neg, \wedge, \lambda, \forall$, a stock of individual variables, a stock of constants, a stock of n-place function, relation and predicate symbols, lambda abstraction and a special belief predicate B. Now let D be our domain of individuals and for each n let P^n be our domain of n-place relations. Then, with respect to a context c and an assignment of individuals to variables g, we have the following:

- If x is a variable, then $[[x]]_{c,g} \in D$;
- If c is a constant, then $[[c]]_{c,g} \in D$;

[7]The other direction does not hold; not all rigid designators are directly referential terms. For example, take a rigidified description. It is by definition a rigid designator, it does however retain its descriptive content

- If R is an n-place relation symbol, then $[[R]]_{c,g} \in P^n$;
- For any context x and assignment y – since these are purely logical symbols – we have that

$$[[\neg]]_{x,y} = NEG,$$
$$[[\wedge]]_{x,y} = CONJ,$$

where NEG and $CONJ$ are functions from truth-values to truth-values.

The variable-binding operations of lambda abstraction and existential quantification are treated with the help of propositional functions. More precisely, the semantic content of $\lambda x Rx, x$, for example, is the function f from individuals to propositions such that if $o \in D$, we have that $f(o)$ attributes the property $R*$ to the pair $\langle o, o \rangle$, or anticipating a bit, $f(o) = \langle \langle o, o \rangle, R* \rangle$. Similarly we can think of an existentially quantified formula $\exists x Rx, x$ as saying that f from above assigns a true propositions to at least one object.

Definition 1 (Semantic Value of Russellian Propositions)**.** With the above comments in mind, we can now see how propositions are assigned to sentences of the language:

Atomic Formulas $[[Pt_1, ..., t_n]]_{c,g} = \langle \langle [[t_1]]_{c,g}, ..., [[t_n]]_{c,g} \rangle, [[P]]_{c,g} \rangle$.

Lambda Abstraction $[[(\lambda x S)t]]_{c,g} = \langle \langle [[t]]_{c,g} \rangle, f \rangle$, where f is a function from individuals to propositions such that $f(o) = [[S]]_{c,g'}$ where g' is an assignment that differs from g at most in setting $[[x]]_{c,g'} = o$.

Conjunction $[[S \wedge Q]]_{c,g} = \langle CONJ, \langle [[S]]_{c,g}, [[Q]]_{c,g} \rangle \rangle$.

Negation $[[\neg S]]_{c,g} = \langle NEG, [[S]]_{c,g} \rangle$.

Existential Quantification $[[\exists x S]]_{c,g} = \langle SOME, f \rangle$, where f is as in the clause for lambda abstraction, and $SOME$ is the property of being a non-empty set.

Belief $[[B(t, S)]]_{c,g} = \langle \langle [[t]]_{c,g}, [[S]]_{c,g} \rangle, BEL \rangle$, where BEL is the belief relation.

But we are not yet done. We have stressed over and again that any satisfactory theory of propositions should attribute to them at least the basic functions that we want propositions to perform. One of the main things we want propositions to be is true or false.[8] For the time being all we have been told is that a proposition is a structured entity made out of objects and properties and the like. Within the possible worlds framework as soon as we are told what a proposition is, we can rest assured that it fulfills its function as a bearer of truth-values. But this is because at the very core of SPW is an equation of sentence meaning to truth-conditions, and semantic values of constituents to intensions. The Russellian does not deny that constituents have extensions and intensions and that sentences thus have compositionally delivered truth values. The difference is that, according to the Russellian, approach there is an additional layer to meaning; as Soames puts it, "[p]ropositional contents do not replace truth-supporting circumstances in a semantic theory; rather, they supplement them with a new kind of semantic value" (Soames 1985, p. 63). But as soon as the Russellian endows propositional constituents with extensions and intentions she is just as able to offer a recursive characterisation of truth as the SPW theorist. What is more, this definition delivers more or less the same truth values a direct reference theorist would assign her sentences within the SPW framework.

[8]For an argument that ordered n-tuples just cannot be things that are true or false refer to the next section.

Extensions and intensions are assigned to objects, properties and propositional functions in the usual way, just as extensions and intensions are assigned to directly referential terms, predicates and function symbols in the possible worlds framework.[9] We may now turn to the definition of truth of a sentence relative to a circumstance, where the o_i's are objects from our domain of objects, P^* is a property from our domain of properties, g is a one-place propositional function, and S^*, Q^* propositions:

Definition 2 (Truth)**.** 1. $\langle\langle o_1, ..., o_n\rangle, P^*\rangle$ is true relative to a truth-supporting circumstance C if and only if $\langle o_1, ..., o_n\rangle$ belongs to the extension of P^* with respect to C.

2. $\langle\langle o\rangle, g\rangle$ is true relative to C if and only if o belongs to the extension of g in C.

3. $\langle CONJ, \langle S^*, Q^*\rangle\rangle$ is true relative to C if and only if S^* is true in C and Q^* is true in C.

4. $\langle NEG, S^*\rangle$ is true in C if and only if S^* is not true in C.

5. $\langle\langle o, S^*\rangle, BEL\rangle$ is true relative to C if and only if o believes S^* in C.

We have now seen the main tenets of the Russellian account. In sections 2.3 and 2.4 we return to the details here as we apply the framework to a variety of examples, examine how it fairs and discuss criticisms that have been raised against it. Before that, however, we return to a question we raised and left unaddressed, namely, what is this relation that figures in, lends structure to and holds together the constituents of Russellian propositions?

2.2.2 King on Structure

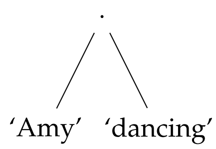
Russellians hold that propositions are ordered n-tuples, or concatenations thereof. Take the simple sentence 'Amy dances'. Following Soames's suggestion, the proposition this sentence expresses is ⟨Amy, dancing⟩, which is but an ordered pair consisting of Amy and the property of dancing. In the previous section we acquired sufficient working knowledge of the Russellian account, and we now know what proposition a given sentence purportedly expresses and how to judge its truth. But we should not be misled by familiarity; as soon as we focus on the metaphysical thesis Russellians are putting forth – identifying propositions with ordered n-tuples – many questions arise. How come 'Amy dances' expresses $\langle Amy, dancing\rangle$ rather than $\langle dancing, Amy\rangle$? How come some n-tuples such as $\langle Amy, dancing\rangle$ are propositions while others such as $\langle 1, 23\rangle$ are not? How is it that some but not all ordered n-tuples come to be carriers of truth values and modal properties? How is it that ordered n-tuples get to be carriers of truth values to begin with? According to King all these, and a few other issues, find answers as soon as we successfully address the problem(s) of the unity of propositions – a problem first noted by the Russell of *Principia*: given that Russellian propositions have constituents – unlike SPW propositions, for example – the question arises of what it is that binds these constituents together or, in other words, what it is that Soames's angle brackets abbreviate. Given we want propositions to – somehow – represent the world as being a certain way, and given that, presumably, Amy and dancing represent Amy and dancing by instantiation, the weigh of carrying this vital representational function falls onto the propositional structure of $\langle Amy, dancing\rangle$. But according to King, such a role cannot be fulfilled by a mere ordering relation, for take Amy, and then take the property of dancing – it is not at all obvious the compound represents the world as one in which Amy dances.

[9] Note that according to SPW a term t or predicate P is assigned an intension, while here we have that the semantic values of terms and predicates, say object o and property P^* – which are *not* their PW intensions – are assigned PW intensions.

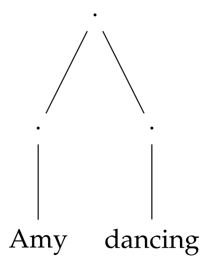
King has developed a very sophisticated account of propositional structure under which all problems of unity disappear – or so it is claimed. We turn to examine his account and the way it addressed questions of unity, discuss how and how successfully concerns such as the ones opening the current sections are addressed, and conclude with a general discussion of the resulting theory.

To begin with, the later King (King 2007) is motivated by the idea that propositions are not the sort of things that have truth values in their own right, independently of minds and languages. Instead, King describes his account as a *naturalized* account of propositions, for he holds that it is us speakers of a natural language who endow propositions with truth conditions. This motivating concern lies at the heart of his account of propositional structure as we see shortly.

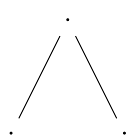
For King propositions are not n-tuples or concatenations thereof; n-tuples and concatenations thereof may be convenient shorthands, but really what binds propositional constituents together is much more complex of a relation than could be hoped to be captured by the angle brackets. If we take the Russellian n-tuple $\langle$Amy, dancing$\rangle$, then we have Amy and we have the property of dancing in an ordered pair. We want propositions to represent features of the world. How Amy and dancing represent features of the world is clear: they represent by instantiation. What remains to answer is what the, *qua* King, complex relation that binds these propositional constituents together and provides the proposition with structure is and how it represents. To answer the latter question we first turn to LF and trees. The sentence 'Amy dances' has the following structure in LF:



Following King, we may extend this structure to the following:



The following part of this latter structure



is the very same relation that the lexical items 'Amy' and 'dancing' stand in in the LF structure of the sentence 'Amy dances'. The lines/ portions of the relation extending the nodes to Amy and dancing in the latter structure are just the semantic relations the words 'Amy' and 'dances' bear to Amy and dancing. Therefore, King concludes, the relation that Amy and dancing stand in in the structured proposition is the relation that is the result of composing the sentential relation as represented in the first structure with the semantic relations the lexical items bear to their semantic values. To be precise, this is a two-place relation, for King existentially quantifies over lexical items as follows: there are lexical items a, b that have as their semantic values Amy and dancing respectively and occur in an LF structure whose sentential relation is as in the first figure.[10]

This is good by way of introduction but does not quite do justice to the later, naturalised King. We need to know how this complex relation represents, and this is where the 'naturalised' part of King's account comes in. King holds that the reason the sentence 'Amy dances' is true if and only if Amy dances is because we speakers *interpret* the sentential relation of the first figure as *ascribing* the semantic value of 'dances' to the semantic value of 'Amy'. It is the fact that this very relation encodes ascription that makes the English sentence have truth conditions instead of being a mere list of words, and this very fact that is responsible for $\langle \text{Amy, dancing} \rangle$ not being just like any other ordered pair but one that abbreviates a proposition that represents the relevant state of affairs. All we need to do to regiment King's earlier definition of the two-place relation Amy and dancing stand in is add that the relation of the first figure encodes ascription.

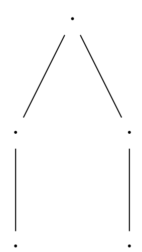
To restate, for King, the complex relation that holds together n propositional constituents in some given proposition is the n-place relation that (a) encodes ascription and (b) that we get by existentially quantifying over the significant lexical items of the sentence expressing the proposition, and composing the sentential relation these lexical items stand in with the relations each item bears to their semantic value. As a result, the proposition in question is the *fact* that such and such semantic values stand in the n-place relation described above. We have now seen how King answers the first unity question of what it is that holds propositions together: he has identified a complex relation in which propositional constituents stand and has reduced the unity of the proposition problem to the problem of the unity of facts, since, according to the current view, propositions are kinds of facts – but certainly not assimilated to the facts which would make the propositions true. As King notes, this reduction is sufficient for "anyone who believes that things stand in relations and possess properties must face the question, if only to dismiss it, of what holds an object and a property together when the object possesses the property or what holds an n-place relation and n objects together when the objects are so related, etc " (King 2013, p.76).

Addressing the second unity question that has been identified – the one pertaining to truth conditions – is straightforward enough for the naturalised King: according to King it is something speakers do that endows the proposition that Amy dances with its truth conditions. In particular, speakers interpret the complex relation Amy and dancing stand in as ascribing the property of dancing located at the relation's right node to Amy at its left node such that the proposition is true if and only if Amy possesses the property of dancing. This is already built in the definition of a proposition *qua* fact, for we have already required that the propositional relation encodes ascription – as per (a). This is then King's Russellian account: propositional constituents stand in complex relations encoding ascription and composed out of sentential relations and semantic relations in which lexical items and their Russellian denotations stand. A proposition is the fact that its constituents stand in such a relation. Propositions do not have truth values independent of us; it is we who endow propositions with truth values by interpreting the complex relation propositional constituents stand in as ascription.

How does this account of propositional unity help us answer the questions with which we

[10] It is important to note as Hanks does (Hanks 2009, p.480) that King's account need not be combined with Russellianism's take on the semantic values of constituteents – although for the most part it does (at least in King 1995) – but that it could be ascribed to by such theorists as we examine in 2.3.2 and 2.3.3.

opened the current section? Our first couple of concerns were hinting at a Benacerraf-style objection against Russellianism: there seem to be no principled way of choosing between equally good reductions of propositions to ordered n-tuples – Soames has chosen $\langle$Amy, dancing$\rangle$ but this is arbitrary; he might as well have gone for $\langle$dancing, Amy$\rangle$.[11] For King the worry phrased us such is no concern at all: King simply does not identify propositions with ordered n-tuples. Granted, the angle bracket representation is still a convenient shorthand, but for King it is no more than a (misleading) shorthand; the relation the angle brackets abbreviate is a highly complex relation as discussed above. However, we may still push the point for we may now doubt whether there is a principled reason to, on the level of syntax, opt for a parsing that assigns 'Amy' to the left node and 'dancing' to the right node and not the other way round, for note that this parsing carries on to propositional structure by means of the sentential relation the latter encompasses. Perhaps this is a question better left to linguists and grammarians. But we may push the point in a different way since King motivates his account as one that does not identify propositions with mathematical objects upon considering the Benacerraf-style objection facing Russellians. And we must grant that, indeed, his account does not identify propositions with ordered tuples but central to his view is that the relation holding propositional constituents together is the composition of the sentential relation binding sentential constituents with the semantic relations each constituent bears to its denotation. Taking 'Amy dances' again, the relation in question is, *qua* King, the following:



But surely this is a mathematical object! It is, as Harry Deutsch exclaims, a "graph(s) of a certain kind (as in "graph theory")" (Deutsch 2008). Of course, King could still hold that the diagram represents the propositional relation in question, but he does not do so and even if he did the main question he purports to have addressed would reopen: what relation is it that binds propositional constituents together?

Next recall our latter couple of concerns: how come some but not all ordered tuples are propositions? How come some of them have truth values while others do not? Again, King does not identify propositions with tuples. However, we may accordingly rephrase: take the relation Amy and dancing stand in in the propositions 'Amy dances', and relate Amy and Bob with it. The result should not be proposition – intuitively it is not meaningful – and King's account can predict as much: it is not as if we ever utter 'Amy Bob', but even when faced with 'Amy Bob' we simply do not interpret the relation relating Amy and Bob as ascription or as anything else that can possibly yield a meaningful content. And lastly, we asked why certain structures have truth conditions while others of the same type do not. This is easy of King; it is precisely because we endow some but not all structures of this type with truth conditions by interpreting their relations certain ways.

King puts forward a detailed and sophisticated account that has many virtues. If adopted, it renders it uncontroversial that propositions exist, it manages to assimilate propositions to facts

[11] For a more detailed discussion see Hanks 2009

(not their truth making facts) and hence reduce problems of unity, it provides a naturalised account of how propositions come to have truth values. However, the account is not without fault. First of all, as discussed above, King does not altogether dispense with Benacerraf-style concerns. However, for the purposes of the current project, King's failure to fully address Benacerraf worries is not a decisive factor. The current is a comparative project and given possible world and extended possible world semantics fall prey to Benacerraf objections too – propositions can equivalently be taken as sets of worlds or characteristic functions thereof – the result cannot be of any help deciding between these alternative accounts of propositions. However, it must be noted that one of King's main motivations behind the current account was that he took the Benacerraf objection against Russellianism to be a serious problem, and to fall prey to a revenge does undermine if not the project then the motivation thereof. But there are further worries, too.

Propositions are according to King ontologically dependent on language: the complex propositional relations he postulates bind propositional constituents together are existential quantifications over lexical items, and composed out if of syntactic and semantic relations. Given this nature of propositional relations, it becomes clear that propositions are language dependent, and given propositions are the bearers of truth and falsity it would seem that truth and falsity are language dependent too. King has defended the latter consequence of his account in King 2007 and perhaps there is a way of making peace with it, as in arguing that truth and falsity existed before and exist independent of minds and language although the bearers of truth and falsity do not. Perhaps King is right in spirit, perhaps a highly naturalised account of propositions is what we should be going after, but King binds content to language inextricably and this seems to produce a number of issues. For example, we often coin sentences to express novel thoughts and meanings, but if we follow King, a novel thought or meaning cannot precede the sentence expressing it. What is more, King leaves no space for non-linguistically expressed propositions, but arguably things other than pieces of language, such as pictorial representations or purely perceptually conceived contents, express propositions. Or to put the point slightly differently, it seems bad to rule out *a priori* as King does that cognitive agents who do not have a language cannot entertain propositions; even animals are often said to entertain propositions! Or take Yablo who has written an entire book (Yablo 2014a) trying to come up with a way to model contents that cannot be expressed by any given sentence. These are contents that Yablo should have no access to whatsoever if King is to be believed. There are all in all strong reasons to hold that propositions are not wholly language dependent, contrary to King.

2.3 Frege's Puzzle and Belief Reports

Now that the theoretical cornerstones of the Russellian account have been presented, we turn to examining certain consequences of the view. Most of these, pertaining to our motivating concerns from chapter 1, are discussed in section 2.4 but amongst them is one that has been addressed extensively and independently by structuralists. This is no other that *Frege's Puzzle* and variations thereof. I have decided to devote a separate section to the issue because I could not do justice to the relevant literature that has been generated by structuralists otherwise. In §2.3.1 we discuss the Russellian strategy of addressing the puzzle, while in the following two sections we examine in turn Crimmins's and Chalmer's alternative structuralist accounts. These can be seen as views arising from dissatisfaction with the Russellian 'way out', while nevertheless remaining structuralist accounts of propositions.

Recall the following pairs of sentences from chapter 1:

$$\text{All woodchucks are whistle-pigs.} \tag{1.5}$$
$$\text{All woodchucks are woodchucks.} \tag{1.6}$$
$$\text{Hesperus is Phosphorus.} \tag{1.7}$$
$$\text{Hesperus is Hesperus.} \tag{1.8}$$

According to the structuralist, following Soames's suggestion from §2.2.1, they express the following propositions respectively:

$$\langle ALL, f\rangle \tag{2.1}$$
$$\langle ALL, f'\rangle \tag{2.2}$$
$$\langle\langle[[\text{Hesperus}]], [[\text{Phosphorus}]], \rangle[[=]]\rangle \tag{2.3}$$
$$\langle\langle[[\text{Hesperus}]], [[\text{Hesperus}]], \rangle[[=]]\rangle \tag{2.4}$$

where f, f' are functions from individuals to propositions such that

$$f(o) = [[\text{woodchuck}(o) \to \text{whistle-pig}(o)]] = \langle\langle\langle o, [[\text{woodchuck}]]\ \rangle, \langle o, [[\text{whistle-pig}]]\rangle\rangle, THEN\rangle$$
$$f(o) = [[\text{woodchuck}(o) \to \text{woodchuck}(o)]] = \langle\langle\langle o, [[\text{woodchuck}]]\ \rangle, \langle o, [[\text{woodchuck}]]\rangle\rangle, THEN\rangle$$

where $THEN$ is a function from truth-values to truth-values that respects the truth-table for material implication, and ALL is a higher-order function that "says" of f, f' that they map all objects to true propositions.

Now if $[[\text{Hesperus}]] = [[\text{Phosporus}]]$ and $[[\text{woodchuck}]] = [[\text{whistle-pig}]]$, according to the structuralist (2.1), (2.2) express the same proposition, and (2.3), (2.4) express the same proposition. But these equations are *prima facie* problematic. For example, (2.1), (2.3) may express informative propositions, while (2.2), (2.4) cannot, and the latter seem knowable *a priori* while the former not. Such differences are manifested as differences in truth-value as soon as we embed the sentences in intentional contexts:

$$\text{Tama fears that (2.1).} \tag{2.5}$$
$$\text{Tama fears that (2.2).} \tag{2.6}$$
$$\text{The ancients believed that (2.3).} \tag{2.7}$$
$$\text{The ancients believed that (2.4).} \tag{2.8}$$

If one holds a relational view of intentional attitudes, such that *belief* and *fear* express a relation between a subject and a proposition, then we have the following regimentation of (2.5) to (2.8):

$$\langle\langle[[\text{Tama}]], (2.1)\rangle, FEAR\rangle \tag{2.9}$$
$$\langle\langle[[\text{Tama}]], (2.2)\rangle, FEAR\rangle \tag{2.10}$$
$$\langle\langle[[\text{The ancients}]], (2.3)\rangle, BEL\rangle \tag{2.11}$$
$$\langle\langle[[\text{The ancients}]], (2.4)\rangle, BEL\rangle \tag{2.12}$$

But since (2.1), (2.2) express the same proposition, and (2.3), (2.4) express the same proposition, (2.9) = (2.9) and (2.11) = (2.12), and their truth conditions are as follows: (2.9) = (2.10) is true if and only if [[Tama]] stands in the relation of fear to (2.1) = (2.2), and (2.11) = (2.12) is true if and only if [[The ancients]] stand in the relation of belief to (2.3) = (2.4).

But strong intuitions dictate otherwise. In the case of Hesperus and Phosphorus, it seems reasonable to suppose that prior to the substantial astronomical discovery that they are both Venus, the ancients did not believe that Hesperus is Phosphorus, although they certainly believed the analytic truth that Hesperus is Hesperus. In the case of Tama, we may quote Ripley:

... [T]his is the wrong prediction. Suppose Tama is familiar with both woodchucks and whistle-pigs, but isn't sure that they are the same kind of critter. He's noticed the similarities, though, and so he has his suspicions. Suppose further that Tama knows he is allergic to whistle-pigs, and knows that he has just been bitten by a woodchuck. In this scenario, [(2.5)] is likely true, while [(2.6)] is almost certainly false. (Ripley 2012, p.104)

Along the way of deriving these counterintuitive consequences, we have, alongside structuralism, made two further assumptions: first, that $[[\text{woodchuck}]] = [[\text{whistle-pig}]]$, $[[\text{Hesperus}]] = [[\text{Phosphorus}]]$ given that woodchucks are whistle-pigs and Hesperus is Phosphorus, and second, that in attribute reports epistemic and intentional verbs have a purely relational semantics. Russellians agree to both these assumptions and are thus left with the task of explaining away the apparent counterintuitiveness of their predictions (§2.3.1). Other structuralists have denied one or both of the assumptions and we turn to their alternative analyses in §2.3.2 and §2.3.3.

2.3.1 Russellianism

According to Soames (Soames 1987) propositional attitudes have purely relational semantics and proper names are directly referential. Therefore, for Soames, belief or propositional attitude ascriptions that differ only insofar as the sentences expressing them employ different coreferential names, express one and the same proposition and are thus truth-conditionally equivalent. Soames bites the bullet. He even goes as far as expressing surprise that "[m]any seem to think that counterexample to this principle [intersubstitutivity *salva significatore* in propositional attitude contexts] are easy to come by (Soames 1987, p.53)."

Soames explains why we oftentimes intuitively resist such substitutions, invoking the semantics versus pragmatics distinction. Firstly, Soames argues that sentential attitudes are not always a reliable guide to propositional attitudes. To be more precise, assenting or dissenting, uttering or not uttering a sentence does not always tell whether an agent believes or disbelieves the proposition expressed by the sentence. Indeed, as Soames acknowledges, a speaker may assent to a sentence $S(t)$ while dissenting or withholding assent from $S(t')$ which differs from the former only in the substitution of some occurrences of a directly referential term t by the coreferential t'. However, even in such cases, we are not warranted in concluding that the speaker believes $S(t)$ while she does not believe $S(t')$; she must believe either or neither because they are the same proposition.

Take the following example adopted from Soames 1987 as borrowed form Richard: Bob is walking down the street on the phone with Amy. He sees a woman across the street who is followed by a suspicious looking man and believes she is in danger. Unbeknownst to Bob, the woman is Amy. Bob runs across the street to warn the woman but speaks nothing on the phone. If we stopped Bob and quizzed him about his beliefs, he would probably sincerely utter

$$\text{I believe she is in danger} \tag{2.13}$$

pointing to the woman across the street, but not

$$\text{I believe Amy is in danger.} \tag{2.14}$$

Many would now suppose that it is dictated by strong intuition that the former is true while the latter false, were they both to be uttered by Bob in this context. However, Soames argues that the propositions expressed by the above two sentences in Bob's context can be shown to have the same truth-value. Suppose Amy – seeing a man running across the street waving his hands in the air – utters the following over the phone to Bob:

$$\text{The man watching me believes I am in danger.} \tag{2.15}$$

But now

$$\text{The man watching you believes you are in danger.} \tag{2.16}$$

is true taken relative to Bob's context as uttered to Amy. But whether he knows it or not, the following is also true relative to his context:

$$\text{I am the man watching you.} \tag{2.17}$$

But from (2.16) and (2.17) it follows that (2.14) is true relative to Bob's context. According to Soames this strategy can be generalised to substitutional 'puzzles' employing any kind of directly referential term – not necessarily demonstratives – and thus, he concludes, it is demonstrable that propositions expressed by sentences of the form 'A *v*'s that S' and 'A *v*'s that S*' where *v* is an intentional verb and S, S* differ only with respect to some directly referential term having been substituted for a coreferential one, have the same truth-conditions. But it also shows how assent is not a reliable guide to belief, for Bob does believe the proposition expressed by (2.14) but would not utter it or assent to it. But the main explanation is pending: why is it that our intuitions push against this as strongly as in Frege's puzzle and the woodchucks–whistle-pigs examples?

According to Soames propositional attitude ascriptions report relations between agents and propositions, while they conversationally imply corresponding relations to sentences which usually hold true. For example that 'A *v*'s S' reports that A stands in the relation of *v*-ing to the proposition that S, but it also suggests that, if asked, A would assent to or even utter 'S'. And while this is usually the case, purported counterexamples to the intersubstitutivity of coreferential directly referential terms in propositional attitude ascriptions, are nothing but cases where the conversational implicature is false. What Soames essentially argues is that 'The ancients believed that Hesperus is Phosphorus' is indeed true, but we have been fooled at thinking it is false because indeed the ancients would not assent to and did not utter 'Hesperus is Phosphorus'.

I find this largely unconvincing. First of all, Soames's suggestions seems to get my Fregean intuitions plainly wrong: when I say no, it is not the case the ancients believed that Hesperus is Phosphorus – even having read Soames – I do not mean to convey that they did believe the proposition but would not phrase it as such, I do not wish to talk about the ancients' linguistic practices, I wish to talk about their astronomy and their corresponding belief system. But perhaps more importantly, it seems to me that Soames tries to bite a very large bullet by implicitly appealing to a nowadays controversial assumption underlying the distinction between semantics and pragmatics. The boundaries between semantics and pragmatics are at best blurry, but be that as it may, and let us concede to Soames that it is indeed some pragmatic feature of names and sentences embedding them that lies at the heart of Frege's puzzle, that is, let us say that indeed it is pragmatically imparted information that differentiates between 'Hesperus is Phosphorus' and 'Hesperus is Hesperus'. My question is, so what if the information is *pragmatically* imparted? The question remains, ought we capture it on the level of content or not? For it is no longer believed that as soon as something falls under pragmatics it is automatically to be shut out from a theory of content. Not only would it take a further argument that Soames does not give to reach his 'no', but there is reason to think that, to the contrary, there is a lot going for 'yes'. To begin with, the strength and near unanimity of our Fregean intuitions with respect to substitutivity puzzles suggests just that: even if it is a merely pragmatic difference between 'Hesperus' and 'Phosphorus' that is behind the puzzle, then we had better account for it in our theory of meaning, for otherwise we are left with a substantial set of counterintuitive consequences and are forced to take a revisionist stand towards our current practices of reporting belief. Moreover, as Ripley (Ripley 2012) argues, further embeddings prove sensitive to the intersubstitutivity of coreferential names, and such spread and systematicity further point towards a theory of meaning that compositionally accounts for the purportedly pragmatic features in question.

Soames claims that 'Hesperus' and 'Phosphorus' differ in pragmatics and that is why 'Hesperus is Phosphorus' and 'Hesperus is Hesperus' or embeddings thereof appear to diverge in meaning and truth value. I argue that even if it is a pragmatic feature of the use of different coreferential terms that is behind our intuitions of divergence in meaning and truth-value, it is does not follow that our theory should exclude such features and explain our intuitions away; instead, our intuitions are too strong and such phenomena too widespread and systematic for a satisfactory theory of meaning not to account for these allegedly pragmatic features compositionally and on the level of content. Or, at least, to push for such a theory would take a stronger case. Salmon 1986 is an extensive defence of structuralism against the objection from Frege's puzzle and here I attempt but a summary of the main points he makes. Salmon takes a different approach than Soames does, but one that also weighs on the semantics versus pragmatics distinction; in his own words:

> It is precisely the seemingly trivial premise that '$a = b$' is informative whereas '$a = a$' is not informative that should be challenged, and a proper appreciation for the distinction between semantically encoded and pragmatically imparted information points that way. (Salmon 1986, p.78)

For example, Salmon concedes that an utterance of '$a = b$' imparts information more valuable that of '$a = a$', as for example the linguistic information that 'a' and 'b' are coreferential, but argues that – contra the early Frege – such information cannot rightfully be said to be the meaning of '$a = b$'. Instead it is argued to be merely pragmatically imparted information.[12] Salmon generalises and claims that as soon as all pragmatically imparted information is stripped away, one sees that the meaning of the sentences – the properly semantic content that is left bare – is one and the same. But it takes an argument to get there.

To offer such, via the intersubstitutivity of co-informational propositions in propositional attitude contexts, Salmon moves to belief and knowledge reports. Indeed, in very much the same vane as Soames, Salmon agrees that we often utter or assent to sentences such as 'The ancients did not know that Hesperus is Phosphorus', and that we do so not wishing to talk elliptically or figuratively but truthfully. However, Salmon claims that we fail to speak truthfully; strictly speaking the ancients knew that Hesperus is Phosphorus. But this clashes sharply with ordinary usage; Salmon needs to explain why we so often, unintentionally spread and believe lies of this sort. Purportedly, it is because we systematically confuse the 'that'-clause with something richer, with the semantic content of the clause, *plus* the information it pragmatically contributes. And this systematic tendency stems from the fact that oftentimes a certain mode of acquaintance with an object – echoing Russellian epistemology – is involved when agents entertain singular propositions. This mode of acquaintance is part of the means by which one

[12] The argument here would be that identity statements can be informative and appear to impart information that is not limited to our linguistic conventions; that Hesperus is Phosphorus seems to be an astronomical fact, of astronomical importance, due to an empirical, astronomical discovery and not merely a lesson in English signs for astronomers and non-astronomers alike. This could be challenged but for now I do not dwell; it seems to me reasonable enough.

understands the singular proposition for it is the means by which one is familiar with one of the proposition's constituents – the object in question. According to Salmon, this generates a 'guise' under which singular propositions are apprehended: Venus *qua* the morning star is different from Venus *qua* the evening star, although Venus *simpliciter* is of course self-identical. An utterance of 'The ancients believed that Hesperus is Phosphorus' conversationally implies that the ancients believed of Venus *qua* morning star that it is the same as Venus *qua* evening star, which is false, although strictly speaking the ancients did believe the purely semantic information of 'Hesperus is Phosphorus' which is that Venus *simpliciter* is Venus *simpliciter*. On the other hand, 'The ancients believed that Hesperus is Hesperus' does not imply that the ancients believed the proposition under different guises of Venus for the first and second occurrence of the term referring to the star, and there is no pragmatically implied information that tricks us into proclaiming the proposition true under this latter guise.

To summarise, Salmon believes in representational guises of objects, and that different representational guises may be pragmatically implied by the use of distinct yet coreferential, directly referential terms. Compositionally, the guise may consequently be said to take 'scope' over entire singular propositions, and that is why we are often fooled into believing that sentences employing coreferential terms express different propositions. When such propositions are embedded into propositional attitude contexts of the sort 'A v's that S', we are tempted to understand the proposition that S under the guise we speculate A apprehends the proposition under. But guises belong to pragmatics, and strictly speaking, the semantic content of sentences differing only in their employment of one coreferential, directly referential term instead of another is one and the same. Again, Salmon is trying to bite a very big bullet. He remains noncommittal about the nature of guises and appeals to our intuitive grasp of the notion as generated by cases in which we understand things other than propositions under some guise – for I do not think that we have an intuitive understanding of what a propositional guise is. But, contrary to Soames, Salmon has come up with an account that postulates yet *systematically* explains a failure on our behalf to recognise certain singular propositions for the propositions that they are, and with a *systematic* account of what the pragmatically imparted information that obfuscates things is. What is more, his account does not seem to disrespect pre-theoretic Fregean intuitions – if anything, his solution comes rather close to Frege's own: guises are hugely reminiscent of Fregean senses, but for the fact that senses are pushed into pragmatics and stripped of their powers to determine semantic content.

Phrased as such this is of course no small difference, but in a review of Salmon's book, Forbes (Forbes 1987) argues that Salmon's account really is just a notational variant of Fregeanism. Forbes's strategy to back up this rather strong claim is to construct a regimentation of Salmon's thesis under Fregean while respecting the main insights of the original. Following Forbes (Forbes 1987), we represent the main thesis of Salmon's account by the following formula:

$$\text{A believes that } p \text{ iff } (\exists x)[\text{A grasps } p \text{ by means of } x \wedge BEL(\text{A}, p, x)], \tag{S}$$

where A is an agent, p a proposition, x ranges over guises under which A is acquainted with p, and BEL is a three place relation holding between an agent, a proposition and a propositional guise just in case the agent believes the proposition under that guise. What is more, we may write $f(\text{A,S})$ to pick out the guise under which the agent A is acquainted with the proposition expressed by the sentence S given that S is used to express the proposition at hand. Now, following Forbes, we consider a version of Fregeanism that allows the referent of a proposition to be a state of affairs – a Russellian proposition effectively – replace "thinks of" for "grasps", regard p as a variable over states of affairs, x as a variable over propositions and define $BEL(\text{A}, p, x)$ as $B(\text{A}, x) \wedge p = \mathit{Ref}(x)$, where B is a relation that holds between agents and (Fregean) propositions they believe, and the right hand side of (S) defines "p is believed by A to obtain":

$$p \text{ is believed by A to obtain iff } (\exists x)[\text{A thinks of } p \text{ by means of } x \wedge B(\text{A}, x) \wedge p = \mathit{Ref}(x)]. \quad \text{(F)}$$

This is plausibly enough a translation a Fregean would perform on Salmon's thesis, and a thesis that she herself might adopt: the Fregean views Salmon's propositional guises as propositions, Salmon's propositions as the referents of her Fregean propositions, and like sense determines reference, here too we have that referents of Fregean propositions – states of affairs – are determined mediately, via guises.

Branquinho (Branquinho 1990) doubts that Forbes's regimentation proves the dissimilation of Salmon's account into a variant of Fregeanism, his point being that (S) and (F) do not make the same predictions therefore they are not equivalent principles. Consider an agent A who, based on her visual experience with Venus at dawn thinks that Phosphorus is a star, and based on her visual experience with Venus at night thinks that Venus is not a star but a planet. Then, in accordance with Salmon's account, we get the following:

$$BEL[\text{A}, \langle \text{Venus, being a star} \rangle, f(\text{A, 'Phosphorus is a star'})] \tag{2.18}$$
$$BEL[\text{A}, \langle \text{Venus, not being a star} \rangle, f(\text{A, 'Hesperus is a star'})], \tag{2.19}$$

which, by (S) give us respectively:

$$\text{A believes that Venus is a star} \tag{2.20}$$
$$\text{A believes that Venus is not a star.} \tag{2.21}$$

Therefore – and flag this point for later – Salmon's account predicts that A has contradictory beliefs. Whether this is a desirable conclusion or not, for now Branquinho's point is that Forbes's Fregean regimentation fails to deliver the same result; (F) does not predict that A has contradictory beliefs. To see that let $[[p]]$ abbreviate the Fregean proposition that p, and read angle brackets as states of affairs – referents of Fregean propositions and not propositions themselves as Salmon would have it. Given our story we have the following:

$$\text{A thinks of } \langle \text{Venus, being a star} \rangle \text{ by means of [[Phosphorus is a star]]}$$
$$\wedge B(\text{A, [[Phosphorus is a star]] }) \wedge \langle \text{Venus, being a star} \rangle = Ref(\text{[[Phosphorus is a star]]}) \tag{2.22}$$
$$\text{A thinks of } \langle \text{Venus, not being a star} \rangle \text{ by means of [[Hesperus is not a star]]}$$
$$\wedge B(\text{A, [[Hesperus is not a star]] }) \wedge \langle \text{Venus, not being a star} \rangle = Ref(\text{[[Hesperus is not a star]]}). \tag{2.23}$$

Therefore, by (F) we have that:

$$\langle \text{Venus, being a star} \rangle \text{ is believed by A to obtain.} \tag{2.24}$$
$$\langle \text{Venus, not being a star} \rangle \text{ is believed by A to obtain.} \tag{2.25}$$

Do (2.24) and (2.25) attribute contradictory belies to A? According to Branquinho they do not. To do so they would have to either amount to or entail

$$\text{A believes } \langle \text{Venus, being a star} \rangle \text{ obtains,} \tag{2.26}$$
$$\text{A believes } \langle \text{Venus, not being a star} \rangle \text{ obtains} \tag{2.27}$$

respectively. Are we to read (2.24) and (2.25) *de dicto*, as in (2.26) and (2.27)? Presumably not for it is vital for Forbes's purposes to keep the states of affairs outside the scope of belief so that they refer to their usual referent and not to their customary senses – it would otherwise seem that his mapping between states of affairs and Russellian propositions, and guises and Fregean propositions crumbles. Does a *de re* reading of (2.24) and (2.25) entail the *de dicto* one in (2.26) and (2.27)? It is widely assumed that not while *de dicto* entails *de re* the opposite does not hold. Therefore, Branquinho concludes, Forbes's Fregean regimentation fails to attribute contradictory beliefs to A, and therefore proves non-equivalent to Salmon's thesis.

Granted, Forbes's thesis was perhaps too strong; it seems that guises are not disguised Fregean senses, or at least Forbes's regimentation fails to prove they are. However, there is an undeniable similarity between Salmon's theory and Fregeanism: Russellian propositions are understood via guises, propositional guises being the result of composing the guises under which the constituents of the proposition are apprehended. But contrary to Fregean senses, Salmon insists that guises are semantically inert and do not determine what proposition is ultimately expressed. And there are two ways to criticise this. Firstly, one may dislike the resulting account, for it attributes contradictory beliefs where we might not see any, it entails that our common practices of belief ascription are largely false, and that we are often unaware of what propositions we express, assent to or utter. This is again the 'too large a bullet'-argument. But moreover, we might question whether Salmon has given us good enough reasons to bite the bullet. To be sure, *Frege's Puzzle* raises a powerful criticism against Fregeanism and other alternatives to Russellianism, but that the theory is all-other-things-being-equal preferable is irrelevant to whether the way he chooses to explain away our intuitions is satisfactory. To motivate the latter, Salmon like Soames, attempts to convince us that our intuitions stem from purely *pragmatic* grounds. And again, even if this is so, I find there is a premise missing. So what if guises belong to pragmatics? Why does that mean that they are inert on the level of content? And what exaggerates frustration is that Salmon's account is already so systematic that a theory of meaning incorporating guises into semantics is but a step away: Salmon is a variant of Fregeanism as soon as guises are allowed to determine propositional content. Locking away guises may keep us safe from the dangers of Fregeanism, but that they are creatures of the pragmatics realm is not a damning conviction in and of itself.

Perhaps there is further reason, perhaps one has a more traditional view on the role pragmatics, perhaps Salmon's otherwise appealing account can be saved. But it is not quite there yet, and no matter, it remains a fact that Russellians are asking us to swallow some very counterintuitive consequences their accounts predict. They are telling us that Frege's puzzle is not really a puzzle, that we often say falsities, that ultimately we should adopt a revisionist stance towards our practices of belief ascriptions.[13] Many – and among them many structuralists, too – simply do not buy that 'Hesperus is Phosphorus' and 'Hesperus is Hesperus' are one and the same proposition and that propositional attitude reports embedding them have the same truth conditions and modal properties. Moreover, given structuralist accounts were motivated by the failure of SPW to capture hyperintensionality, it may seem rather disappointing to settle for biting the bullet; one would expect a hyperintensional semantics to deliver fine-grained enough content to distinguish between necessary equivalents that strong intuitions dictate differ cognitively or truth-conditionally. We now turn to structuralist accounts that attempt just that: they deny one or more of the Russellian assumptions and built theories that can accommodate our strong intuitions concerning variants of Frege's puzzle.

[13] Bealer 2004 goes as far as to argue that approaches such as Salmon's and Soames's that appeal to pragmatics in order to avoid the issues posed by Frege's Puzzle to direct reference reveal an inconsistency in the theories.

2.3.2 Crimmins and Perry

Crimmins and Perry are direct reference theorists but – contra the Russellians – deny that *belief* or *knowledge* have purely relational semantics. However, they do hold a closely related yet independent thesis, one that has been dubbed *semantic innocence*: sentences embedded in belief contexts express the same proposition they would express if they were not embedded. Assuming *direct reference*, *semantic innocence* and all clauses of Soames's definitions of structured propositions and their truth other than the ones pertaining got belief, they go on to built a theory of beliefs and belief reports that accommodates our Fregean intuitions towards the latter; that is, a theory that predicts that belief reports embedding sentences that differ only in their use of coreferential directly referential terms may express different propositions.

Let us take an example and stick with it throughout:

Superman wears a blue cape.
Clark Kent wears a blue cape.

According to the current view these two sentences express the following propositions respectively – treating 'wearing a blue cape' as a property for simplicity:

$$\langle[[\text{Superman}]], [[\text{wearing a blue cape}]]\rangle \tag{2.28}$$
$$\langle[[\text{Clark Kent}]], [[\text{wearing a blue cape}]]\rangle \tag{2.29}$$

and, moreover, assuming direct reference, $[[\text{Superman}]] = [[\text{Clark Kent}]]$ therefore the propositions (2.28) and (2.29) are one and the same. So, assuming *semantic innocence* as they do, how is it that the theorists get out of the Russellian predicament?

According to Crimmins and Perry, beliefs are concrete cognitive structures, particulars that belong to an agent and related to the world in a way that allows us to classify them by propositional content (Crimmins and Perry 1989, p.688). Therefore, the authors allow themselves to speak of particular beliefs – as mental tokens – instead of limiting themselves to talk of belief relations between a person and a proposition. This is where they give up purely relational semantics for beliefs. The details of their alternative account follow, and while here they are presented merely as theses, the authors do spent a considerable amount of space motivating their assumptions.

Beliefs are structured entities – structured much like Russellian propositions – but, given they are cognitive particulars, they contain cognitive particulars as constituents. More precisely, they contain 'ideas' and 'notions' where these are terms of art: loosely speaking, 'ideas' are ways of thinking about properties and 'notions' are ways of thinking about objects, the properties or objects ideas or notions are ideas or notions of being their content. Thus, beliefs acquire a proposition as content. To take an example, agent A's belief that Superman wears a blue cape is the ordered pair consisting of A's idea of wearing a blue cape and A's notion of Superman in that order. Thus, the proposition believed by A is the ordered pair consisting of the content of A's idea of wearing a blue cape and the content of A's notion of Superman, in that order.

If we buy this, then it is easy to account for failures of intersubstitutivity of co-referential terms in belief reports. Let us take agent A and the pair of sentences from above. If indeed A believes that (2.28) but A does not believe that (2.29), then this can be explained on the basis of A employing different notions of Superman and Clark Kent. Let us represent A's notion of Superman by S, A's notion of Clark Kent by C and A's idea of wearing a blue cape by B. Then, we have that a has a belief $\langle S, B\rangle$ whose content is $\langle$Superman, wearing a blue cape$\rangle$. If A does not believe that Clark Kent wears a blue cape, then A has no belief whose content is $\langle$Clark Kent, wearing a blue cape$\rangle$, no belief $\langle C, B\rangle$. In particular, $\langle S, B\rangle$ is clearly distinct from

$\langle C, B \rangle$, and its content is different from $\langle$Clark Kent, wearing a blue cape$\rangle$ precisely because S is not a notion of Superman, regardless of Superman being identical to Clark Kent.

The current should suffice to convince us that Crimmins and Perry do indeed manage to avoid the counter-intuitive consequences Frege's puzzle forced upon Russellians while remaining structuralists, direct reference theorists and semantically innocent. Whether we should buy their theory of beliefs is a big question and one for which we should consult with cognitive scientist. I attempt no such thing. Instead allow me to flag two points: first, it is not Crimmins's and Perry's *structuralism* that gets them out of trouble; indeed their theory of beliefs could equally as well be paired with SPW and help the latter deal with Frege's puzzle. And as Ripley notes, if the strategy works,

> ... [t]hen, it undermines half of the fineness-of-grain argument for structuralism; coarse-grained propositions should suffice for attitude reports. (Ripley 2012, 107)

Secondly, Crimmin's and Perry's solution is customed to address a very specific class of problems; this strategy allows structuralism to draw hyperintensional distinctions only in the cases where necessary equivalents are embedded in epistemic contexts. By design, the theory has nothing to say when such propositions are embedded, for example, in conditionals. More on those later in §2.4.

2.3.3 Chalmers

Chalmers is first of all a two-dimensionalist; his preferred semantics associates with expressions in contexts two kinds of intensions, a *primary* and a *secondary* one. Intensions are, as per usual, functions from possibilities to extensions, and for Chalmers, there are two distinct sets of possibilities that are relevant to meaning. Firstly, associated with secondary intensions, is the usual set of metaphysically possible worlds, rendering Chalmers's secondary intensions of simple expressions equivalent to their usual SPW meanings. Secondly, those which Chalmers calls *scenarios* give rise to primary intensions. Scenarios are standardly identified with centered worlds, that is, ordered triples consisting of a world, an individual and time within the world, where the latter pair is seen as the 'centre' of the scenario. To get a rough idea of how primary intensions work, let us look first at expression types that are parts of sentences – for Chalmers's account is compositional in both intensions. To take an example, following Chalmers, " [t]he primary intension of a paradigmatic use of 'Hesperus' may function, very roughly, to pick out a bright object visible at a certain point in the evening sky in the environment of the individual at the center of the scenario." (Chalmers 2011, p. 5). Roughly speaking, Chalmers allows that we think "of primary intensions as a sort of descriptive content associated with an utterance of an expression" (Chalmers 2011, p.5). Moving up to complete sentences, we say that the primary intension of a sentence S is true at scenario c just in case, if the subject at the centre of c were to know they were inhabiting that particular scenario c, then they would accept S, which hints at the epistemic flavour primary intensions carry.

Thus far, Chalmers has not made clear what he takes propositions to be and it is a matter of choice what single entity to identify a proposition with within this two-dimensional framework. Chalmers wants his propositions to cut finely and do as much explanatory work as needed. Given the dichotomy between primary and secondary intensions, Chalmers has managed to divorce the carriers of modal properties and those of epistemic ones: secondary intensions of sentences typically reflect modal properties while secondary ones typically reflect epistemic properties, as suggested by the following example. Take 'Hesperus is Phosphorous'. In a post-Kripkean spirit, it is rather uncontroversial that the sentence expresses a metaphysically necessary proposition, and looking at its secondary intension, the latter is true in all possible worlds: just as in the SPW framework, as soon as 'Hesperus' and 'Phosphorus' are to be taken directly referential, rigid designators, the secondary intension of 'Hesperus is Phosphorus' does not fail in any possible world. However, it is largely agreed upon that the sentence is not *a priori* – a significant astronomical discovery was involved in the establishment of its truth – and this epistemological status it holds is reflected in the primary intension of the sentence: 'Hesperus' and 'Phosphorus' may have distinct descriptive contents for agents in scenarios – think of emphthe morning start versus *the evening star* – therefore there are scenarios in which the primary intension of the sentence 'Hesperus is Phosphorus' is false. Surely modal and epistemic properties of propositions are both important and thus Chalmers wishes to accommodate for both primary and secondary intensions in his notion of a proposition. But, moreover, noting that "the explanatory benefits of allowing logical form in propositions are well known" (Chalmers 2011, p.6), Chalmers moves on to defend a structured account of propositions, and that is why we examine his theory within the current chapter.

Given this two-dimensional framework and his wish to incorporate logical form by means of structuralism, the first step Chalmers takes towards his account of propositional content is a passing from primary and secondary intensions of expressions *simpliciter*, to structured ones. This is simple. Chalmers's structured secondary intensions are just Russellian propositions, and to construct structured primary intensions is no more technically involved: the structured primary intension of a sentence is the structure consisting of the primary intensions of all the constituents of the sentence, as the sentence appears in LF, and structured according to the sentence's logical form. At this point, Chalmers could hold that a proposition is a pair of structures, a structured primary intension, and a structured secondary one. However, it is not precisely this pair Chalmers identifies propositions with, at least not in Chalmers 2011. Instead, he introduces the notion of an *enriched intension*, the enriched intension of a simple expression being an ordered pair of the expression's primary intension and its possible-worlds extension, and the enriched intension of a sentence being the structure built out of its constituents's enriched intensions compositionally and structured according to the sentence's logical form. It is then enriched intensions of sentences that Chalmers identifies propositions with. Let us take the example from above and consider an utterance of 'Hesperus is Phosphorus'. The enriched intension of 'Hesperus' is going to be the pair (H, V), where H is the primary intension of 'Hesperus' and V is the star Venus, and similarly for 'Phosphorus' (P, V), and for identity $(I, =)$. Then the sentence expresses the structured proposition $\langle\langle(H, V), (P, V)\rangle, (I, =)\rangle$.

According to Chalmers enriched propositions are more natural alternatives than the pairing of primary and secondary structured intensions, and already suffice for most of the explanatory roles propositions are meant to play. Moreover, primary intensions of sentences and Russellian propositions corresponding are recoverable from enriched propositions in an obvious manner, the primary intension of 'Hesperus is Phosphorus' being $\langle\langle H, P\rangle, I\rangle$ and its secondary extension/ Russellian proposition $\langle\langle V, V\rangle, =\rangle$, while evaluating the secondary extension across different possible worlds we recover the structured secondary intension of the sentence. We then say that an enriched proposition is true at a world if its secondary intension is, and true at a scenario if its primary intension is. Enriched propositions are necessary if they are true at all worlds and *a priori* if they are true at al scenarios, thus offering a natural explanation of Kripke's necessary *a posteriori* and contingent *a priori* statements. Moreover, enriched propositions can behave in a largely Fregean manner. Primary intensions in particular seem to play part of the roles of Fregean senses, as for example in that *a priori* inequivalent expressions such as some coreferential names have distinct primary intensions. All in all, Chalmers's propositions cut very finely. Given they are structured entities, they can distinguish between any necessary equivalents that other structuralist can distinguish between – such as distinct mathematical truths – but owing to their Fregean/ descriptivist flavour, they can draw even more hyperintensional distinctions such as between propositions expressed by sentences differing only in their embedding distinct coreferential terms.

Being able to draw the latter distinction – as between 'Hesperus is Phosphorus' and 'Hesperus is Hesperus' – is quite naturally central to Chalmers's success at dealing with Frege's puzzle, although Chalmers in his account of propositional attitude ascriptions does not retain a purely relational semantics for belief. More on Chalmers on belief shortly, but for now, I wish to underline an aspect of Chalmers's theory which, although not immediately obvious, is of great significance. Chalmers can distinguish between the propositions expressed by 'Hesperus is Phosphorus' and 'Hesperus is Hesperus' on the basis of the propositions's primary intensions, owing to the primary intensions of 'Hesperus' and 'Phosphorus'; that is, 'Hesperus' and 'Phosphorus' have distinct primary intensions. When introducing the theory, we said we can think of primary intensions as some sort of descriptive content associated with expressions but that does not mean that Chalmers appeals to a descriptivist semantics for names. To the contrary, it seems like Chalmers is a direct reference theorist what with his endorsement of Russellian propositions. Granted, Russellian propositions pertain to secondary intensions but it would be highly dubious if a two-dimensionalist appealed to different semantics for expression types for each of her dimensions. How is then that, assuming names are directly referential, Chalmers can still attribute distinct primary intensions to'Hesperus' and 'Phosphorus'? It is because the worlds that are allowed to take up the first place of the scenario triplets are not limited to the possible ones. More precisely, the primary intensions of 'Hesperus' and 'Phosphorus' are allowed to differ, only because there are scenarios in which Hesperus is not Phosphorus. Perhaps primary intensions communicate some sort of descriptive content, but on the bottom line they are functions from possibilities to extensions, extensions of names are objects, and only in so far as in some scenario the objects the names refer to differ do we get the much desired divergence in primary intension. Flag this point: Chalmers appeals to impossible worlds, and it is in virtue of appealing to impossible worlds and not to propositional structure that he can face Frege's puzzle. We now turn shortly to his account of belief ascriptions for we have underlined over and again that Chalmers is a structuralist that denies both assumptions of Russellianism: plain relational semantics for belief and purely Russellian semantics for names.

Chalmers much in the way of Crimmins, assumes that beliefs are structured mental representations.[14] However, for Chalmers they are mental representations that can be associated with both primary and secondary intensions, the primary intension of a belief being the set of scenarios such that if the agent accepts she is finding herself in that scenario then she accepts the belief, and secondary intensions corresponding to the usual way beliefs are treated within the SPW framework. Then, in the same spirit as truth of enriched propositions, the truth conditions of belief ascriptions are given by the following principle:

(APR) 'x believes that *S*' is true of *i* iff *i* has a belief with the Russellian content of *S* (in the mouth of the ascriber) and with an *S*-appropriate structured primary intension (Chalmers 2011, p. 14).

It is left unspecified exactly what makes for an *S*-appropriate structured primary intension, but a couple of things may be noted. For example, for the ancients to believe that Hesperus is the evening star, their belief would have to single out Hesperus with a 'Hesperus'-appropriate

[14]Strictly speaking, Chalmers uses this as a simplifying assumption and insists that no such entities as structured mental representations need be postulated as long as it is allowed for subjects to stand in psychological relations to appropriate structured entities. Not much hinges on this for our current analysis.

primary intension, which might suggest any of a few descriptive contents such as being the evening star, or appearing at a given spot in the night sky, but not some content that is associated with Venus *qua* Phosphorus. That is to say, many an intensions may be S-appropriate but not all.

Still the account is somewhat incomplete for we have not examined what the referents of 'that' clauses are. Although Chalmers does consider alternative accounts, he eventually endorses an account according to which referents of 'that' clause are enriched propositions, and plays around with a mostly relational semantics for belief so as to yield in accordance with (APR). So assuming the referent of 'that S' is the enriched proposition p, we have that 'x believes that S' is true of i just in case i endorses a proposition that is coordinate with p, where endorsing an enriched proposition is roughly having a belief with the same structured content. Note that albeit relational, Chalmers's account poses a further constraint on the semantics of belief that of coordination. Coordination is meant to relax the appropriate relation that need hold between 'that Hesperus is the evening star' in my mouth and the proposition the ancients endorsed, so as to rule out cases in which an attitude ascription may be true while the one the attitude is ascribed to may not endorse precisely that proposition.[15] This theory in place, it is clear that Chalmers can make the right predictions for (2.7), (2.8): the primary intensions of 'Hesperus' and 'Phosphorus' differ and although the ancients did believe the necessary, *a priori* truth that Hesperus is Hesperus, they did not believe the astronomically involved truth that Hesperus is Phosphorus for they were in possession of two distinct primary intensions for Venus.

Chalmers's theory is very rich but almost a conglomeration of disparate theories: two-dimensionalism, structuralism, impossible worlds semantics. This is not a conclusive criticism, but such a loaded ideology is rather unattractive. If any of two-dimentionalism, structuralism or impossible worlds semantics is able to deliver on its own as much hyperintensional content and has equal explanatory power, then on grounds of economy and simplicity it is to be preferred. What is more, Chalmers's approach at points verges on *ad hoc*; for example, although he starts from an intensional account of meaning – and a two-dimensional one for that – he swiftly moves to structured content because structured content delivers explanatory power he wants his resulting theory to posses. Again, none of this is decisive and merely because the accounts he composes are traditionally looked upon as rivals should not preclude their composition; what I find unappealing is that Chalmers's composition seems purely additive, his resulting account does not deliver any more than the sum of its theoretical sums does. Let me just conclude that if any of 'purer' structuralist or impossible-worlds account fairs equally as well with the data the Chalmers's account loses to them by default.

But more importantly it is not as if his account is immune; it does, for example, suffer from problems that we noted face structuralism. It is as vital for Chalmers to address the question of what binds propositional content together, and given that propositional content is even richer in Chalmers's view – we have both secondary extensions and primary intensions figuring in the nodes of the structures – an explanation would be, if anything, more cumbersome. However, the main issue that Chalmers's approach raises is one brought out in Ripley 2012 and Jago 2014. The latter theorists argue as follows: the main motivation behind adopting structuralism is to deliver finess of grain, as for example is considered to be required by Frege's puzzle. However, both Crimmins and Chalmers eventually deliver that finess of grain not owing to their structuralists theories *structuralism* but to whatever supplementation they further endorse. In the case of Chalmers this is clearly his appeal to impossible worlds semantics. And given in

[15] For example, I may correctly say that Pierre believes London is pretty, although Pierre would not endorse the proposition expressed by 'London is pretty'.

the current thesis we wish to judge structuralism and impossible worlds semantics mainly according to how well they fair at drawing hyperintensional distinctions, Chalmers's structuralist approach undermines the motivation behind adopting structuralism in the first place, and hence indirectly argues against structuralism but also directly in favour of impossible worlds semantics in his success at dealing with Frege's puzzle.

2.4 Conclusion

2.4.1 The Data and the Verdict

All variants of structuralism we have looked at can distinguish between necessary equivalent propositions so long as the sentences expressing them do merely differ in their embedding distil coreferential terms. This is so because structured propositions have constituents and can be individuated in virtue of the constituents that make them up. For example, although all of (1.1), (1.2), (1.3), (1.4) have the same intension – the set of all possible worlds – they are distinct propositions for $2, 3, e, \vee$ are constituents of each of (1.1), (1.2), (1.3), (1.4) respectively but of no other proposition in this set of four. As a result, all structuralists can distinguish between any embedding of such propositions, be it in propositional attitude reports or in other contexts, simply because the propositions embedded are distinguished to begin with. The exact same reasoning should convince one that structuralists can distinguish between necessary falsehoods that do not differ merely in their use of distinct coreferential terms, and any embedding thereof.

But as we have seen in §2.3, things are not as straightforward when we turn to necessarily equivalent propositions the sentences expressing which differ only in so far as they employ distinct coreferential names. In the aforementioned section, we examined three different approaches to variants of Frege's puzzle. Salmon and Soames each bite the bullet, concede that such propositions are indeed identical and attempt to explain away strong intuitions pulling towards the opposite conclusion. Crimmins and Perry abandon a purely relational semantics for belief, while Chalmers distinguishes between the propositions expressed by, say, 'Hesperus is Hesperus' and 'Hesperus is Phosphorus' by appealing to the primary intensions of the terms employed. Let us examine these options in reverse order.

Chalmers owes the success of his strategy to primary intensions alone and not to structuralism. To see this, consider the following toy theory: a proposition P expressed by sentence S is a pair consisting of the primary and secondary intension of S but neither the primary nor the secondary intension is taken to be structured, they are just sentence intensions in the image of the SPW framework. On this account, the primary intensions and hence the propositions expressed by 'Hesperus is Hesperus' and 'Hesperus is Phosphorus' differ, for there are scenarios in which the former is true while the latter false. This shows that it is not *structuralism* that delivers the finess of grain required to distinguish propositions generating Frege's puzzle. Ripley (Ripley 2012) and Jago (Jago 2014) have taken this to greatly undermine the structuralist project, for it is indeed the promise of hyperintensional content that mainly motivates structured accounts of propositions. Chalmers's strategy is a solution to Frege's puzzle but it is one that argues for extended possible world semantics if anything – for as we have seen, it precisely because impossible worlds figure in scenarios and at such worlds the extensions of necessary propositions may differ.

Crimmin's and Perry's solution is cut to measure for propositional attitude ascriptions and in particular belief reports. As such, we may repeat the argument from above: the mere need for such a supplementation undermines the finess-of-grain arguments put forth for structuralism. But more importantly, it is a strategy that fails as soon as we turn to different embeddings of the relevant kind of propositions. For example, we find in Perry and Crimmins an explanation as to why the proposition expressed by 'Lois Lane believes that Clark Kent is Clark Kent' is true while 'Lois Lane believes that Clark Kent is Supermant' is false. But add to the story that Lois Lane is really afraid of aliens and thus of Superman, then Crimmins and Perry cannot explain why 'Lois Lane fears that Clark Kent is Clark Kent' comes across as false while 'Lois Lane fears that Clark Kent is Superman' appears to be true. However, it is reasonable to suppose that the theorists' account can be uniformly extended to cover intentional and epistemic contexts other than ones generated by belief. However, following Ripley (Ripley 2012) and Jago (Jago 2014), it is not only intentional contexts that prove problematic. Conditional for example seem to generate Fregean intuitions.

Consider the following:

$$\text{If Clark Kent is Clark Kent, then Clark Kent is Clark Kent.} \tag{2.30}$$

$$\text{If Clark Kent is Clark Kent, then Clark Kent is Superman.} \tag{2.31}$$

Intuitively, (2.30) and (2.31) seem to express different propositions. The Fregean intuition that 'Clark Kent is Clark Kent' and 'Clark Kent is Superman' are distinct carries over, so to speak, in these embeddings where we find the propositions expressed by these sentences in antecedent and consequent of indicative conditionals. (2.30) seems uninformative while (2.31) not so, and we can imagine Lois Lane agreeing to the former but not the latter. We may even embed the propositions expressed by the sentences further, in intensional contexts, in which case we see that 'Lois Lane believes that (2.30)' intuitively comes out as true while'Lois Lane believes that (2.31)' does not. The point here is that if one wants to accommodate for intuitions generated by Frege's puzzle on a semantic level, then she is obliged to do the same for the current class of cases. And, as Crimmins's and Perry's theory stands, it falls short of doing so. But one may now argue that indicative conditionals are commonly treated as epistemic, and if thus that a strategy such as Crimmins's and Perry's can be extended to accommodate for this latter sort of data. However, Jago (Jago 2014) argues that a similar case can be made with subjunctive conditionals. Take the following:

$$\text{If Clark Kent were not Superman, then Superman would not be Superman.} \tag{2.32}$$

$$\text{If Clark Kent were not Clark Kent, then Clark Kent would not be Clark Kent.} \tag{2.33}$$

Odd as these may sound, the point Jago is driving is that (2.33) is of the form $P \rightarrow P$, and as such comes out true on any account of $\rightarrow$. What is more, intuitively, while a strange thing to assert, (2.32) does sound true. However, this does not seem to be the case for (2.32); Jago holds that the latter is "*prima facie* false" (Jago 2014, p. 77), and I admit to sharing this intuition. Pre-theoretically, trying to repress familiarity with direct reference theory and rigid designation, it seems intuitively true to me that even if Clark Kent were not Superman he would still be self-identical. Perhaps the intuition can be fought against but I believe these are worries enough to abandon Crimmins's and Perry's strategy which is essentially a supplement for structuralism meant to account for embeddings of propositions in intentional contexts. As such the theory raises two worries: first, that if supplementation is needed to deliver hyperintensionality then the main motivation behind structuralism is severely weakened, rendering the account self-defeating, and secondly that there arguably are non-intentional contexts that generate similar puzzles which the theory is not suited for addressing.

I believe that Soames's and Salmon's position is the one doing the most justice to structuralism. Granted, as we have discussed extensively, the theorists bite a very large bullet and moreover endorse some controversial views in trying to explain away the counter-intuitiveness their stance generates. However, and contrary to what Ripley (Ripley 2012) and Jago (Jago

2014) imply, these structuralists advance no supplementation to structuralism in order to address Frege's puzzle; to the contrary, they do not agree that Frege's is a semantic puzzle and do not want their theory of content to distinguish between propositions expressed by sentences that differ only in so far as they employ distinct coreferential terms. What they supplement their views with are theories to explain away intuitions and make their views palatable. But in explaining away the counterintuitive consequences of their accounts they do not undermine the motivation behind their semantic theory: sure enough, they are Russellians because they want fine content, but if we are to believe them, then there is no distinction to be made between 'Hesperus is Hesperus' and 'Hesperus is Phosphorus'. For this reason, I believe that it is Soames and Salmon that best defend structuralism against the challenges arising from Frege's puzzle. This does not mean that I endorse their view. What it means is that when at a later stage I shall be comparing the relative virtues and vices of the different hyperintensional accounts of propositions presented in this thesis, I will be taking Soames's and Salmon's approach to be the best structuralism has to offer with respect to Frege's puzzle.

One last thing that ought to be mentioned is that structured accounts of propositions have often been criticised for delivering too fine a contents. So far I have been talking as if there is no upward limit on how finely it is desirable that propositions be individuated, but just as having propositions that are too coarse is problematic, so is having propositions that are too fine. To take a rather striking example, if we follow King, then sentences of different languages that intuitively 'say the same thing' do not express the same proposition, for recall that we quantify over lexical items to get the propositional relation that according to King figures in structured contents. This is cutting too fine and is rather problematic. Again, we have certain functions we want propositions to perform – one of them being explaining how distinct sentences may express the same thing – and an account of the nature of propositions that renders them incapable of performing such functions is an unlovely account of propositions. Chalmerian propositions too may be thought to cut too finely. The full account of why this is so required a discussion that I cannot afford to fully indulge into, but according to Chalmers 'I am thirsty' said by me, and 'You are thirsty' said by you about me do not express the same proposition for the primary intensions of the pronouns differ. And Salmon and Soames are not immune to this kind of worry either: according to the most basic of Russellian assumptions, '$2 + 1 = 3$' and '$1 + 2 = 3$' express distinct propositions. What seems like an difficult balance for structuralists to strike is give an account of propositional structure that is constrained by syntax – so to go hyperintensional – but not allow syntax too much power.

The overarching aim of the thesis is to identify the most promising account of propositions on the market, on the assumption that delivering fine grained enough content is of vital importance. In the current chapter I turned to examine variants of structured propositions accounts hoping to find a theory that is an overall satisfactory and just the right amount of hyperintensional. In §2.4.1 – which importantly builds on the discussion in all previous sections – I concluded that the most satisfying variant of structuralism – Russellianism *à la* Soames and Salmon – fairs well with all the hyperintensional data from Chapter 1 but refuses to recognise the need to distinguish between intensionally equivalent propositions expressed by sentences that differ only in so far as they employ distinct coreferential terms. Adopting this stance is asking us to explain away some rather strong Fregean intuitions, and the explanations set forth by Soames and Salmon are not entirely uncontroversial. To add, it is arguable that the theory cuts too finely at points. And lastly, we ought to keep in mind that problems of the unity of propositions are at large when considering structuralist approaches and I have found that the most promising attempt to address such issues, King's from §2.2.2, is partly self-defeating – in that it takes avoiding Benacerraf-style worries to be a main motivation for adopting the account, yet is faced with 'revenge' Benacerraf concerns – and delivers many a counterintuitive consequences stemming from the assumption embedded in the account that propositions are ontologically dependent on language. All in all, structuralism is faced with a fair deal of problems, certainly enough to warrant us to turn to alternative analyses.

Chapter 3

Extended Possible World Semantics

3.1 Introduction – Outline

The current chapter looks into accounts of content that extend the SPW framework by introducing impossible worlds. Loosely speaking, impossible worlds are 'ways the world could not have been', just as possible worlds are 'ways the world could have been'. We are not concerned here with a restricted notion of impossibility, for the totality of possible worlds is not delivered to us by a restricted notion of possibility either: the totality of possible worlds is not the totality of worlds that are physically possible or possible given Newtonian mechanics; rather it is the totality of worlds that are possible in the widest, most absolute sense. Metaphysical, logical and perhaps mathematical modalities are commonly considered to take widest scope, and thus, dually, our talk of impossible worlds concerns worlds that are metaphysically, logically, or mathematically impossible. But there are conceptions of impossible worlds other than our 'ways' talk, which – although aiming to capture absolute impossibility –nevertheless deliver a subset of the totality discussed above. To take some, impossible worlds can alternatively be considered as worlds where the laws of one's favourite logic fail (paraconsistent, paracomplete and relevant logics making appearances), as worlds where the laws of most people's favourite logic – classical logic – fail, as worlds where most people's favourite law of their favourite logic fails – the Law of Non-Contradiction. These different conceptions deliver different totalities of impossible worlds, and these totalities are ordered by inclusion, correspondingly to the order of the above exposition. I do not take a stance on which conception, if any, is better. Each extended possible worlds semantics we look at comes with its own assumptions, and I treat the current issue purely heuristically: the theory that delivers the best results on the hyperintensionality front is the one whose conception of impossible worlds I adopt.

Another issue I remain quiet about is the motivation behind the introduction of impossible worlds. Just as is the case for possible worlds, the strongest arguments in their favour come from their heuristic value.[1] If by the end of this chapter we are convinced that impossible worlds are valuable for delivering hyperintensionality – as I hope the reader will be – then this should suffice for current purposes. Relatedly, I am not going to focus on questions pertaining to the metaphysics of impossible worlds. A short discussion of some basic and related issues can be found in §3.4, but even there the focus is on convincing the reader that metaphysical issues can and should be black-boxed given the overall aim of my thesis.

As we shall see, upon extending the SPW framework to include impossible worlds, the way

[1] For a discussion see Berto 2013.

we model content and epistemic-doxastic states of agents is very much in the spirit of SPW semantics. To give an indication, propositions are still identified with sets of worlds – including but not limited to possible words – and knowledge and belief are still modelled as special modal operators. However, from the very outset the strategy seems promising if we keep an eye on hyperintensionality, for the worlds we are dealing with now are more and not constrained by logical possibility and closure.

Here is then how the current chapter is laid out. I begin with an exposition of Priest's (Priest 2005) impossible world semantics for a first order modal language with intentional operators. For expository purposes, I first present his semantics for a language without identity, because the motivating concerns behind the framework form quite an independent class – Priest is trying to overcome logical omniscience and closure under entailment for intentional operators. Identity is then added to the language and the semantics are extended, while looking into how the framework deals with problematic cases stemming from interchanging co-referential names in intentional contexts. Priest succeeds in destroying logical omniscience and closure under entailment for intentional operators, but Jago (Jago 2014) argues he goes too far. A good account of rationality should recognise that certain truths are trivial, certain falsehoods obvious, certain inferences under intentional operators worth preserving. With this being his main motivating concern, Jago goes on to construct a highly sophisticated framework that purportedly strikes just the right balance between the idealised notion of rationality promoted by the SPW approach and the highly illogical one advanced by Priest. The chapter concludes with a a comprehensive outlook on how these two extended possible worlds frameworks fare with the hyperintensional data from §1. Jago's account is argued preferable, not only on the basis of its treatment of the data but in its internal coherence, too. However, I find that Jago's account, albeit delivering almost all the finess of grain I set out to find, falls short, which I attribute to the framework's failure to properly capture *aboutness* features. Finally, an afterword on some related metaphysical issues – so that I say more than nothing on the matter – brings the chapter to a close.

3.2 Priest

In the current section I present the world semantics found in Priest 2005, both the formal framework and various philosophical motivations behind it. Following Priest, I present the semantics for a first order language without identity ($\mathcal{L}$) first, and consequently show how to extend the framework to interpret a language with identity ($\mathcal{L}^+$). I have decided to break the exposition into two for exegetical purposes: as will shortly become clear, the semantics for $\mathcal{L}$ are constructed in an attempt to overcome a particular class of issues arising from the treatment of intentionality within standard possible world semantics, such as logical omniscience, and closure under entailment, while those for $\mathcal{L}^+$ help us deal with an additional and somewhat independent class of issues, such as Frege's puzzle. So on we go!

3.2.1 Semantics for $\mathcal{L}$

The basis of $\mathcal{L}$ is a first-order modal language without identity. That is, the logical vocabulary of $\mathcal{L}$ comprises of $\neg, \vee, \wedge, \rightarrow, \forall, \exists, \Box, \Diamond$ and the non-logical vocabulary of $\mathcal{L}$ includes a stock of variables, constant symbols, n-ary predicates and n-ary function symbols.[2] But we now extend

[2] Priest does not use these quantifier symbols in Priest 2005. The reason for doing so is that he is a *noneist*; loosely speaking, he holds that there are objects that do not exist. The existential quantifier, through a long standing tradition

the language by adding a collection of intentional operator symbols, which, following Priest, we symbolise by upper case Greek letters. Anticipating a bit, if t is a term and A is a formula, then $t\Psi A$ is a formula that we may read as 't Ψ's that A'.

An interpretation I for $\mathcal{L}$ is a tuple $\langle \mathcal{P}, \mathcal{I}, \mathcal{O}, @, D, \delta \rangle$. $\mathcal{P}$ is the set of possible worlds and $\mathcal{I}$ the set of impossible worlds. These two sets are disjoint and together they form the set of *closed* worlds, that is $\mathcal{P} \cup \mathcal{I} = \mathcal{C}$. $\mathcal{O}$ is the set of *open* worlds, it is disjoined from $\mathcal{C}$ and together they form the totality of worlds: $\mathcal{C} \cup \mathcal{O} = \mathcal{W}$. $@ \in \mathcal{P}$ is the actual world, and D is the domain of objects, and δ assigns every non-logical symbol a denotation as follows:

$$\text{If } c \text{ is a constant, then } \delta(c) \in D; \tag{3.1}$$

$$\text{If } f \text{ is an } n\text{-place function symbol, then } \delta(f) \text{ is an } n\text{-place function on } D; \tag{3.2}$$

$$\begin{aligned}&\text{If } P \text{ is an } n\text{-place predicate and } w \in W\text{, then } \delta(P, w) \text{ is a pair } \langle \delta^+(P, w), \delta^-(P, w) \rangle,\\ &\text{where } \delta^+(P, w), \delta^-(P, w) \in D^n;\end{aligned} \tag{3.3}$$

$$\begin{aligned}&\text{If } \Psi \text{ is an intentional verb symbol, then } \delta(\Psi) \text{ is a function from } D \text{ to } W \times W\\ &\text{such that } \delta(\Psi)(d) = R^d_\Psi\text{, i.e. a binary relation on } W.\end{aligned} \tag{3.4}$$

The first two clauses should be familiar, but let me expand on the latter two. If P is an n-place predicate symbol, then the interpretation function assigns it both an extension – intuitively the set of n-tuples of which P holds relative to a world w – and an anti-extension – the set of n-tuples of which P does not hold relative to W. If w is a possible world, then we require that for all P, $\delta^+(P, w) \cap \delta^-(P, w) = \varnothing$ and that $\delta^+(P, w) \cup \delta^-(P, w) = D$, that is, we want the extension and anti-extension of all predicates at possible worlds to be mutually exclusive and exhaustive of the domain of objects. These requirements are relaxed at impossible and open worlds, and we get worlds at which an object may be both P and not P or both.[3] Our last clause says that the semantics for intentional operators are basically a generalisation of usual binary semantics for modal operators. If we take a Ψ and an individual d, then its interpretation is a binary relation R^d_Ψ on W such that $(w, w') \in R^d_\Psi$ if and only of thongs in w' are as d Ψ's them to be in w. For particular intentional operators, such as K for knowledge, we may want to impose constrains on the accessibility relation they correspond to – such as reflexivity to account for the idea that knowledge is factive – but this is an issue we do not currently address.

Before looking into how to extend the interpretation function δ and give truth and falsity conditions for formulas, a few exegetical points are due. Firstly, we are dealing with constant domain modal semantics, that is, the domain of objects at each world is one and the same – D. However, Priest does not hold that for any object and any world that object exists at that world. Priest is a Meinongian, or as he would prefer to call it a noneist, in that he holds that there are objects that do not exist. Therefore, even though the domain remains constant across worlds, we do not have the counterintuitive consequence that everything exists necessarily, for what exists at each world is a subset of what lies in the domain of that world.

epitomised by Quine is considered by most as ontologically committing. That is the well-known Quinean motto that 'to be is to be the value of a variable'. However, there is an opposing and long standing tradition, that of Meinongianism, which distinguished between to be and to exist. Being is being the value of a variable, but existence is captured by a real, non-trivial predicate. We cannot afford to get into the details here but let us make this clear: according to Priest the quantifiers are not ontologically committing. Ontological commitment is to be expressed by a non-trivial existence predicate E, a predicate that is *not* defined parasitically on the quantifiers as $Ex := \exists x(x = x)$, but placing a restriction on D and metaphysically 'thickened' in some way or another. I black-box such debates here, and stick to the usual quantifier symbols.

[3] As Priest notes, we have a choice of where and how to relax these conditions. In the current exposition I relax both conditions at impossible and open worlds but not at possible ones which makes the semantics correspond to a relevant logic. If the conditions where relaxed at possible worlds too, then we would have a paraconsistent logic.

Secondly, this is as good a place as any to offer some initial motivation and give some intuition behind $\mathcal{I}$ and $\mathcal{O}$. Priest extends the standard possible world framework to include impossible worlds for two reasons: to avoid fallacies of relevance – so he can have a relevant and not a classical conditional – but also because he holds conditional statements have a closer connection to logic than other complex formulas.[4] To be more precise, Priest argues that a conditional statement of the form $A \rightarrow B$ expresses an entailment, a logical law. Add to that that impossible worlds are meant to be worlds where laws of logic fail, and we get that at impossible worlds we want conditional formulas to behave truth-conditionally differently than in possible ones. The way Priest achieves this result formally is by, essentially, treating conditional statements like atomic formulas at impossible worlds.[5] And moreover, by relaxing conditions of exhaustivity and exclusivity of predicate extensions and anti-extensions at impossible worlds already, we have that at such worlds the Law of Non-Contradiction and Excluded Middle can fail. However, for any world in $\mathcal{I}$ conjunctions, disjunctions, and quantifiers still work as usual. Therefore, if we stopped at $\mathcal{I}$, we would still get closure under entailments pertaining to these connectives – i.e. $A \wedge B \vdash A$ – and agents would still be omniscient as far as most truths not in conditional form are concerned. To destroy closure under entailment and logical omniscience, Priest introduces $\mathcal{O}$ – worlds where all formulas are treated essentially as atomic, hence worlds not closed under any non-trivial consequence – thus 'open' – and with the ability to make any logical truth whatsoever false.

The above should suffice as exegetical notes on some of the choices that become clear in the formal framework we now resume presenting. As per usual, truth and falsity values are assigned relative to an assignment of variables to denotations. So let s be an assignment of objects to variables, then our interpretation function assigns the following denotations to terms:

Definition 3 (Denotations of Terms)**.**

1. For c a constant: $\delta_s(c) = \delta(c) \in D$;
2. For x a variable: $\delta_s(x) = s(x) \in D$;
3. For f a function symbol and $t_1, ..., t_n$ terms: $\delta_s(f(t_1, ..., t_n)) = \delta(f)(\delta_s(t_1), ..., \delta_s(t_n))$.

Before moving on to truth and falsity, we need to introduce a technical notion, that of a *matrix*, that Priest makes a lot of use of since it significantly simplifies the exposition. In Priest's own words: "Suppose that the term t occurs in the formula $A(t)$. Say that t occurs free if it contains no occurrence of a free variable that is bound in $A(t)$. [...] Call a formula a matrix, if all its free terms are variables, no free variable has multiple occurrences and – for the sake of definiteness – the free variables that occur in it, $x_1, ..., x_n$, are the least variables greater than all the variables bound in the formula, in some canonical ordering, in ascending order from left to right." (Priest 2005, p. 17) So, for example, assuming the canonical order is x_0, x_1, x_2, and P is a one-place predicate, Q a two-place predicate and g a one-place function, f a two-place function and t a term, then $Px_1 \rightarrow (\exists x_0)(Pf(x_0, x_2))$ is a matrix but $\forall x_0 Q(x_0, g(x_2), t)$ is not. This is straightforward enough; the thing to stress is that for any formula A it can be obtained by a unique matrix, say $\bar{A}$, by substituting some terms in A for the appropriate free variables, making sure such terms are free in the resulting matrix. Now, the claim is, that for any matrix $C(x_1, ..., x_n)$ of the form $A \rightarrow B$, or $\Box A$, or $\Diamond A$ at impossible worlds and any matrix $C(x_1, ..., x_n)$ of any form whatsoever at open worlds, δ assigns an extension and anti-extension to C such that $\delta^+(C, w), \delta^-(C, w) \in D^n$. We may now finally give truth and falsity conditions for a sentence,

[4] I do not discuss relevant conditionals in the thesis, but for a discussion of the connection between impossible worlds and relevance once can look at Berto 2013.

[5] Modal formulas too are treated as atomic in impossible worlds, and the motivation behind this follows shorty.

at a world, relative to our interpretation $\mathcal{I}$ and our assignment of objects to variables s. For 'A is true at w' or 'w makes A true' we write $w \Vdash_s^+ A$, while for 'A is false at w' or 'w makes A false' we write $w \Vdash_s^- A$.

Definition 4 (Truth and Falsity at Closed worlds)**.** 1. For any $w \in \mathcal{C}$

$$w \Vdash_s^+ P(t_1, ..., t_n) \iff \langle \delta_s(t_1), ..., \delta_s(t_2) \rangle \in \delta^+(P, w)$$
$$w \Vdash_s^- P(t_1, ..., t_n) \iff \langle \delta_s(t_1), ..., \delta_s(t_n) \rangle \in \delta^-(P, w).$$

2. For any $w \in \mathcal{C}$

$$w \Vdash_s^+ A \wedge B \iff w \Vdash_s^+ A \text{ and } w \Vdash_s^+ B$$
$$w \Vdash_s^- A \wedge B \iff w \Vdash_s^- A \text{ or } w \Vdash_s^- B.$$

3. For any $w \in \mathcal{C}$

$$w \Vdash_s^+ A \vee B \iff w \Vdash_s^+ A \text{ or } w \Vdash_s^+ B$$
$$w \Vdash_s^- A \vee B \iff w \Vdash_s^- A \text{ and } w \Vdash_s^- B.$$

4. For any $w \in \mathcal{C}$

$$w \Vdash_s^+ \exists x A \iff \text{for some } d \in D, w \Vdash_{s[x/d]}^+ A$$
$$w \Vdash_s^- \exists x A \iff \text{for all } d \in D, w \Vdash_{s[x/d]}^- A,$$

   where $s[x/d]$ is an assignment of objects to variables that differs from s only insofar as it assigns d to x.

5. For any $w \in \mathcal{C}$

$$w \Vdash_s^+ \forall x A \iff \text{for all } d \in D, w \Vdash_{s[x/d]}^+ A$$
$$w \Vdash_s^- \forall x A \iff \text{for some } d \in D, w \Vdash_{s[x/d]}^- A,$$

   where $s[x/d]$ is an assignment of objects to variables that differs from s only insofar as it assigns d to x.

6. As we discussed above, although we want negation to behave classically at possible worlds, we want impossible worlds to invalidate logical laws such as the law of non-contradiction and the law of excluded middle. That is, we want it to be the case that at some impossible worlds for some formula A it may be the case that both A and $\neg A$ are true, or that neither are. This is taken care of by relaxing the constraints of exclusivity and exhaustivity at impossible worlds, and we need not add anything to our semantic clause:[6]

7. For any $w \in \mathcal{C}$

$$w \Vdash_s^+ \neg A \iff w \Vdash_s^- A$$
$$w \Vdash_s^- \neg A \iff w \Vdash_s^+ A.$$

[6]This is one of the alternatives that Priest discusses but not the account he necessarily endorses. He seems to want to relax constrains in possible worlds too, but here we only do so at impossible and open worlds.

8. As we discussed above, Priest wants conditional statements to behave differently at impossible worlds so that we may capture that at impossible worlds logical entailments fail. This is reflected in the definition of truth and falsity of conditional formulas at impossible worlds. Moreover, Priest wants his conditional to be relevant, which is reflected in the clause for possible worlds – note that in the relevant clause we do not only consider possible but impossible worlds too:

   - For any $w \in \mathcal{P}$

$$w \Vdash_s^+ A \to B \iff \text{for all } w' \in \mathcal{C} \text{ such that } w' \Vdash_s^+ A, w' \Vdash_s^+ B$$
$$w \Vdash_s^- A \to B \iff \text{for some } w' \in \mathcal{C} \text{ such that } w' \Vdash_s^+ A, w' \Vdash_s^- B.$$

   - For any $w \in \mathcal{I}$, let $C(x_1, ..., x_n)$ be any matrix of the form $A \to B$, and let $t_1, ..., t_n$ be terms that can be freely substituted for $x_1, ..., x_n$. Then we have that

$$w \Vdash_s^+ C(t_1, ..., t_n) \iff \langle \delta_s(t_1), ..., \delta_s(t_n) \rangle \in \delta^+(C, w)$$
$$w \Vdash_s^- C(t_1, ..., t_n) \iff \langle \delta_s(t_1), ..., \delta_s(t_n) \rangle \in \delta^-(C, w)$$

9. Modal operators are meant to capture the laws of logic, and what is logically possible or necessary at impossible and open worlds differs from what is so at possible ones. Moreover, impossible and open worlds should be modally inaccessible from possible worlds so to make sure that laws of logic do come out necessary at possible worlds. With that in mind here is what we get:

   (a) - For any $w \in \mathcal{P}$

$$w \Vdash_s^+ \Box A \iff \text{for all } w' \in \mathcal{P}, w' \Vdash_s^+ A$$
$$w \Vdash_s^- \Box A \iff \text{for some } w' \in \mathcal{P}, w' \Vdash_s^- A$$

   - For any $w \in \mathcal{I}$, let $C(x_1, ..., x_n)$ be any matrix of the form $\Box A$, and let $t_1, ..., t_n$ be terms that can be freely substituted for $x_1, ..., x_n$. Then we have that

$$w \Vdash_s^+ C(t_1, ..., t_n) \iff \langle \delta_s(t_1), ..., \delta_s(t_n) \rangle \in \delta^+(C, w)$$
$$w \Vdash_s^- C(t_1, ..., t_n) \iff \langle \delta_s(t_1), ..., \delta_s(t_n) \rangle \in \delta^-(C, w)$$

   (b) - For any $w \in \mathcal{P}$

$$w \Vdash_s^+ \Diamond A \iff \text{for some } w' \in \mathcal{P}, w' \Vdash_s^+ A$$
$$w \Vdash_s^- \Diamond A \iff \text{for all } w' \in \mathcal{P}, w' \Vdash_s^- A$$

   - For any $w \in \mathcal{I}$, let $C(x_1, ..., x_n)$ be any matrix of the form $\Diamond A$, and let $t_1, ..., t_n$ be terms that can be freely substituted for $x_1, ..., x_n$. Then we have that

$$w \Vdash_s^+ C(t_1, ..., t_n) \iff \langle \delta_s(t_1), ..., \delta_s(t_n) \rangle \in \delta^+(C, w)$$
$$w \Vdash_s^- C(t_1, ..., t_n) \iff \langle \delta_s(t_1), ..., \delta_s(t_n) \rangle \in \delta^-(C, w)$$

10. For intentional operators, so to not get unwanted results such as logical closure or logical omniscience, we allow closed worlds to intentionally access open worlds where logical laws may fail and truths are world-distinguished from their logical consequents. That is, we have that, for any $w \in \mathcal{C}$:

$$w \Vdash_s^+ t\Psi A \iff \text{for all } w' \in \mathcal{W} \text{ such that } (w, w') \in R_\Psi^{\delta(t)}, w \Vdash_s^+ A$$
$$w \Vdash_s^- t\Psi A \iff \text{for some } w' \in \mathcal{W} \text{ such that } (w, w') \in R_\Psi^{\delta(t)}, w \Vdash_s^- A.$$

We may now define truth and falsity for formulae at open worlds. These are completely anarchic worlds and we formally impose that by treating any formula whatsoever essentially as atomic, making use of matrices as in some clauses of the above definition.

Definition 5 (Truth and Falsity at Open Worlds)**.** For any atomic formula or formula of the form $\neg A, A \wedge B, A \vee B, A \rightarrow B, \Box A, \Diamond A, t\Psi A$ let $C(x_1, ..., x_n)$ be a matrix of that form, and let $t_1, ..., t_n$ be terms that can be freely substituted for $x_1, .., x_n$ respectively. Then, for any $w \in \mathcal{O}$, we have that:

$$w \Vdash_s^+ C(t_1, ..., t_n) \iff \langle \delta_s(t_1), ..., \delta_s(t_n) \rangle \in \delta^+(C, w)$$
$$w \Vdash_s^- C(t_1, ..., t_n) \iff \langle \delta_s(t_1), ..., \delta_s(t_n) \rangle \in \delta^-(C, w).$$

Although we shall not be concerned with the matter much – the main focus being on the above definitions of truth and falsity – it is good to note that logical consequence is truth preservation at @ in all interpretations, and logical truth is truth at @ in all interpretations. Thus setting things up, where intentional operators are not concerned, we get a fully relevant logic – that is, whenever $\vdash A \rightarrow B$, A and B share a predicate or propositional parameter – but if one defined the conditional classically then the logic corresponding to this semantics would be classical. Lastly, as previously noted, if one allowed that extensions and anti-extensions of predicates at possible worlds too can be overlapping and non-exhaustive, then the logic would be paraconsistent – that is, it we would have $A, \neg A \not\vdash B$, and dually $A \not\vdash B \vee \neg B$.[7]

Back to our main focus, now the semantics is set up, we return to some of the motivating concerns behind this construction, discuss the relevant limitations of possible world semantics for intentionality, and examine how the current framework is supposed to address them. The main issues that Priest is concerned with is, firstly that of logical omniscience. The standard possible world treatment of intentional operators would have it that whenever A is a logical truth so is $t\Psi A$, for some term t and intentional verb Ψ. As discussed in §1, this is a highly counterintuitive result for most intentional verbs. For example take knowledge K, logical omniscience dictates that all agents know all logical truths, but intuitively this is absurd: there are highly complex logical axioms that took years to establish their truth and certainly most non-logicians do not know them. Or take belief and assume that the Principle of Excluded Middle is a logical truth, it thus follows that all intuitionists believe it. Or take fear and that if Fido is a dog, then Fido is a dog, it thus follows that you fear that if Fido is a dog, then Fido is a dog. It is easy to see how the current framework invalidates logical omniscience. Let A be any logical truth. There is an impossible or an open world, say w', such that $w' \Vdash_s^- A$. Take a term t and an intentional operator Ψ at a possible world w. Then we can have it such that $(w, w') \in R_\Psi^{\delta(t)}$, and hence by clause (10) of Definition 4 that $w \Vdash_s^- t\Psi A$, as required.

To give a counterexample, take the Law of Excluded Middle and consider $Pa \vee \neg Pa$. We give a model such that $@ \Vdash_s^- t\Psi(Pa \vee \neg Pa)$:

$$D = \{a\}$$
$$\mathcal{P} = \{@\}$$
$$\mathcal{I} = \{w\}$$
$$R_\Psi^{\delta(t)} = \{(@, w)\}$$
$$\delta^+(P, w) = \delta^-(P, w) = \varnothing$$

[7] Priest 2005 proves these results – or references other works of his where the results are proven – throughout §1.

By definition 4 we have that

$$w \Vdash_s^- Pa$$

and

$$w \Vdash_s^- \neg Pa,$$

therefore

$$w \Vdash_s^- Pa \vee \neg Pa$$

and since $(@, w) \in R_\Psi^{\delta(y)}$, we have that

$$@ \Vdash_s^- t\Psi(Pa \vee \neg Pa).$$

Priest's second main worry concerns closure under entailment with respect to intentionality. The issue here is that according to the standard treatment of intentionality within possible world frameworks, if B is a logical consequence of A, then if agent t Ψ's that A, then agent t Ψ's that B, and this is precisely because if B is a logical consequence of A then B is true at all worlds where A is true. This makes significantly counter-intuitive predictions. To take one, let A be $P \vee \neg P$ and let B be a highly complex logical truth that no one has ever considered. Clearly, B being a logical truth, it is a logical consequence of A. However, although many can be said to know or believe that A, by construction, no one has ever considered, let alone come to believe or know that B. Or as Priest invites us to do, suppose that the Peano Axioms entail Fermat's Last Theorem. One may well believe of each Peano Axiom that it is true, but neither believe nor disbelieve Fermat's Last Theorem. Or consider that I desire to go out for King's Day, but if I go out for King's Day I will not make my deadline. It does not follow that I desire to not make my deadline, to the contrary, I do. Whether this is irrational or not is not the point; the point is that intentionality does not seem to be closed under entailment. Priest's open worlds suffice to destroy closure under entailment for intentional operators, for assume that B is a logical consequence of A. Granted, at all closed worlds where A is true B is also true, but there may be an open world, say w', such that $w' \Vdash_s^+ A, w' \Vdash_s^- B$, and it may be the case that for some possible world w, term t and intentional operator Ψ we have that $(w, w') \in R_\Psi^{\delta(t)}$ and hence, by clause (10) of Definition 4, that $w \Vdash_s^- t\Psi B$.

To see a counterexample, consider the following. Pa is a logical consequence of $Pa \wedge Pb$, but the following is a model such that an agent *Psi*;s that $Pa \wedge Pb$ but not that Pa:

$$\begin{aligned}
&D = \{a, b\}\\
&\mathcal{C} = \{@\}\\
&\mathcal{O} = \{w\}\\
&R_\Psi^{\delta(t)} = \{(@, w)\}\\
&\delta^+(Px, w) = \varnothing\\
&\delta^+(Px \wedge Py, w) = \{(a, b)\}.
\end{aligned}$$

Then, by definition 5, we have that

$$w \Vdash_s^+ Pa \wedge Pb,$$

but

$$w \Vdash_s^- Pa,$$

and hence, by definition 4 and the fact that $(@, w) \in R_\Psi^{\delta(t)}$, we have that

$$@ \Vdash_s^+ t\Psi(Pa \wedge Pb)$$

but

$$A \Vdash_s^- t\Psi Pa$$

as required.

Lastly, although not so commonly noted as the above two concerns, the Barcan and converse Barcan formulas for arbitrary intentional operators are validated by standard possible world treatments of intentionality, but again, such principles are intuitively quite problematic. Take an arbitrary intentional operator Ψ, formula $A(x)$ – with x free – and term t, then the Intentional Barcan formula is $(\forall x)(t\Psi A(x)) \Vdash t\Psi(\forall x)(A(x))$, and the converse Intentional Barcan $t\Psi(\forall x)(A(x)) \Vdash (\forall x)(t\Psi A(x))$. What the Barcan formula 'says' is that if for all x t Ψ's that they are A, then t Ψ's that all x are A. But this is not true in general, for consider that for each thing I may believe that it is P but not believe that all things things are P, for I may not believe that the former are all the things there are. The Converse Barcan 'says' that if an agent Ψ's that for all things that they are A, then for each thing the agent Ψ's that it is A. Although this is a bit more intuitive, it does not apply for arbitrary intentional operators, for consider, following Priest, that I fear everyone does not love me. It may still be the case that I do not fear that Donald Trump does not love me; I could not care less about Donald Trump's affections, what I fear is not that for this person and that person they do not love me, instead what I fear is that everyone may be like that.

The current semantics invalidate the Intentional Barcan formulas. For example we have that $(\forall x)(t\Psi Px)\neg \Vdash t\Psi(\forall x)(Px)$ as is established by the following counter model:

$$\begin{aligned}
&D = \{a\} \\
&\mathcal{C} = \{@\} \\
&\mathcal{O} = \{w\} \\
&R_\Psi^{\delta(t)} = \{(@, w)\} \\
&\delta + (\forall x Px, w) = \varnothing \\
&\delta^+(Px, w) = D
\end{aligned}$$

One can easily check that, according to definitions 4 and 5, we have that $@ \Vdash^+ t\Psi(\forall x)(Px)$ is true, but that $@ \Vdash^+ t\Psi(\forall x)(Px)$ is not true.

Priest is clearly successful in destroying logical omniscience, closure under entailment and invalidating the Intentional Barcan formulas. However, there is the worry that the solution is in a way too cheap, or rather as Priest words it, that the solution "destroys all inferences concerning intentional operators" (Priest 2005, p. 24). Priest seems to dismiss this worry quite lightheartedly, noting that not all such inferences fail, for example various inferences concerning the quantifiers such as $t\Psi Pa \Vdash (\exists x)(t\Psi Px)$ still hold. However, granted even if not all, it seems that one too many an inferences are destroyed. For example, as in the second counter-model constructed above, the framework allows that agents may know a conjunction but not its conjuncts. This is rather counterintuitive, and the only seems to be to somehow distinguish between open worlds, include some but not all. This is a significant concern and we return to this point in the section about Jago, who augments the basic framework aiming to strike a balance between avoiding logical omniscience and closure under entailment while saving some inferences concerning intentional operators. But before that we turn to extending our semantics to interpret a language with identity, and looking at issues arising at the intersection of identity and intentionality. Relatedly, Soames's objection mentioned in §2 is presented, and it is shown how Priest's framework avoids it.

3.2.2 Semantics for $\mathcal{L}^+$

In the previous section we recognised a number of motivating concerns arising from the treatment of intentionality in standard possible world semantics, followed Priest in developing an extended possible world semantics and looked at how these issues may be thus resolved. In the current section we examine another class of issues that arise from the interplay between identity and intentional operators, extend the semantics to interpret a language with identity – following Priest – and discuss how these issues are resolved. We close the section by considering a famous argument raised in Soames 1987, purportedly establishing that any theory that identifies propositions with sets of truth-supporting circumstances and endorses a number of allegedly uncontroversial assumptions makes a validates a certain class of arguments that are intuitively not sound. We then show that Priest's framework actually blocks the argument in question, and thus does not fall prey to Soames's criticism.

In §2 we devoted a considerable amount of space discussing Frege's Puzzle, focusing in particular on propositional attitude reports. The issue was that, intuitively, the propositions expressed by 'The ancients believed that Hesperus is Hesperus' and 'The ancients believed the Hesperus is Phosphorus' differ in truth value – the former being true while the latter false – although the sentences expressing them differ only insofar as they employ distinct yet coreferential rigid designators – 'Hesperus' and'Phosphorus'. To phrase it somewhat differently, the following appears to be an invalid inference:

$$\frac{\begin{array}{l}\text{Hesperus is Phosphorus.}\\ \text{The ancients believed that Hesperus is Hesperus.}\end{array}}{\text{The ancients believed that Hesperus is Phosphorus.}}$$

However, given standard possible world semantics for intentional operators, the above piece of reasoning comes out valid. This is essentially because according to possible world semantics all instances of the Schema of the Substitutivity of Identicals (SI) below – including the above inference – are valid:

SI: $t_1 = t_2, A(t_1) \Vdash A(t_2)$, whenever t_1, t_2 are free when substituted in $A(x)$.

Intentional contexts aside, SI seems almost definitional of what it means for things to be identical; if two terms t_1, t_2 (rigidly) denote the same object o, and given that the object has a certain property, say $P(o)$, if that $P(o)$ can be expressed by $A(t_1)$, then $A(t_1)$ is true, and so is $A(t_2)$, given it expresses the very same thing, namely that $P(o)$. As Priest notes in (Priest 2014, p. 59) "[t]he left-to-right direction quickly delivers SI, given second-order universal instantiation and *modus ponens* ...", where the biconditional concerned is no other than Leibniz's Law as formulated in second order logic: $a = b \iff \forall X(Xa \equiv Xb)$. Bracketing the issue of whether there are contentious instances of SI where intentional contexts are not concerned – as was thought to be the case for modal contexts pre-Kripke, for example – and given that our focus is on intentionality, it is the following schema that we want our new semantics to invalidate:

SII: $t_1 = t_2, t\Psi A(t_1) \Vdash t\Psi A(t_2)$, whenever t_1, t_2 are free when substituted in $A(x)$.

I will be assuming that where intentional contexts are not concerned, SI is a principle we want our semantics to validate, for it seems bad to invalidate a principle that is almost self-evident given that, for all we care here, among instances of SI that are not instances of SII, we do not find any counterexamples. But we want our new semantics to invalidate SII precisely because there are many contentious instances of it such as Frege's Puzzle and The Hooded Man Puzzle Priest discusses in §2 of (Priest 2005), and to which we now turn.

Suppose you are in a room and a hooded man is brought in through the door. Further assume that upon entering the room you baptise the hooded man 'Clark' – an assumption we make to avoid obfuscating the issue at hand by entering discussions of whether definite descriptions are rigid designators or not. Now, unbeknownst to you, Clark is actually your brother Bob, who you very well know suffers from a heart disease. The following inference, and instance of SII, seems invalid:

$$\begin{array}{l}\text{Bob is Clark.}\\ \text{You know that Bob has a heart disease.}\\ \hline \text{You know that Clark has a heart disease.}\end{array}$$

The premises are intuitively true, while the conclusion false for you do not know anything of the hooded man brought into the room, let alone his confidential medical record. Note that it makes no difference to proclaim that, really, you do know that Clark has a heart disease, you just don't realise that he does, for the problem can resurface like so:

$$\begin{array}{l}\text{Bob is Clark.}\\ \text{You realise that Bob has a heart disease.}\\ \hline \text{You realise that Clark has a heart disease.}\end{array}$$

Nor is there anything special about knowledge in this example. The puzzle can be recreated by changing 'know' for 'realise' as above or for 'believe', or even for 'fear'. Change the example slightly, and assume that you do not know that Bob has a heart disease but he has been in and out of the hospital, checking his heart because the doctors are worried. Then, being a loving brother, you do fear he has a heart disease:

$$\begin{array}{l}\text{Bob is Clark.}\\ \text{You fear that Bob has a heart disease.}\\ \hline \text{You fear that Clark has a heart disease.}\end{array}$$

But surely, you do not particularly care about the hooded man walked through the door, and even if you did, you have no reason to suspect, let alone fear, that he has a heart problem.

Priest takes this to show that it is SII that is the problem, not the semantics for descriptions, nor peculiarities of some intentional operators. And assuming, as we have, that we want to save all instances of SI that are not instances of SII, we now present a semantics for a language with identity – to be found in Priest 2005 that meets our goals.

We extend $\mathcal{L}$ by adding a special binary predicate '$=$', which is treated as a logical symbol in that its denotation is the same across all models. The interpretation is a tuple $\langle \mathcal{P}, \mathcal{I}, \mathcal{O}, @, D, Q, \delta \rangle$, where $\mathcal{P}, \mathcal{I}, \mathcal{O}, @$ are as before. Even at this early stage, we need a bit of motivation for explaining Priest's Q and D. Recall Bob and Clark from the Hooded Man Paradox:

> These two might have the same identity; they might not. That is, there are worlds compatible with all that you know in which they do have the same identity, and worlds in which they do not. In particular, then, objects may have different identities at different worlds. For any object, then, there is a function that maps it to its identity at each world. Indeed, for technical simplicity, we can just identify the object with this function. Thus, we can take an object to be a map from worlds to identities. (Priest 2005, p. 43)

Q is then a set of things we may think of as identities, following Priest's suggestion, and D is a collection of functions from W to to Q such that if $d \in D$, then $d(w) \in Q$ is the identity of d at w. The interpretation function δ works as follows:

Definition 6 (Interpretation Function).

- If c is a constant, then $\delta(c) \in D$;
- If f is an n-place function symbol, then $\delta(f)$ is an n-place function on D;
- If P is an n-place predicate, then for all $w \in W$, $\delta(P, w) = \langle \delta^+(P, w), \delta^-(P, w) \rangle$, where $\delta^+(P, w), \delta^-(P, w) \in Q^n$;
- If C is a matrix of the form $A \to B, \Box A, \Diamond A$, then for all $w \in \mathcal{I}, \delta(C, w) = \langle \delta^+(C, w), \delta^-(C, w) \rangle$;
- If C is a matrix of any form, then for all $w \in \mathcal{O}, \delta(C, w) = \langle \delta^+(C, w), \delta^-(C, w) \rangle$;
- For all $w \in \mathcal{P}, \delta(=, w) = \langle \{(o, o) \mid o \in Q\}, Q^2 \setminus \{(o, o) \mid o \in Q\} \rangle$.

Given an assignment of values to variables s, a denotation is assigned to terms very much like before. The difference now is that that denotations are assigned from D, which means that the denotetion of a term is a function from worlds to identities. That being the case, we need to slightly adjust our truth and falsity conditions:

Definition 7 (Truth and Falsity).

1. For any $w \in W$

$$w \Vdash_s^+ P(t_1, ..., t_n) \iff \langle \delta_s(t_1)(w), ..., \delta_s(t_n)(w \rangle \in \delta^+(P, w))$$
$$w \Vdash_s^- P(t_1, ..., t_n) \iff \langle \delta_s(t_1)(w), ..., \delta_s(t_n)(w) \rangle \in \delta^-(P, w)$$

2. If $C(x_1, ..., x_n)$ is a matrix of the form $A \to B, \Box A, \Diamond A$ and $w \in \mathcal{I}$, or if it is a matrix of any sort and $w \in \mathcal{O}$, then if $t_1, ..., t_n$ are terms freely substitutable for $x_1, ..., x_n$ in C:

$$w \Vdash_s^+ C(t_1, ..., t_n) \iff \langle \delta_s(t_1)(w), ..., \delta_s(t_n)(w \rangle \in \delta^+(C, w))$$
$$w \Vdash_s^- C(t_1, ..., t_n) \iff \langle \delta_s(t_1)(w), ..., \delta_s(t_n)(w \rangle \in \delta^-(C, w))$$

3. The rest of the clauses for the connectives, quantifiers, modal and intentional operators remain the same.

The essential difference between the current and the previous semantics is that we now consider objects to be functions from worlds to identities. This may struck some as odd, especially since in this post-Kripkean era it is largely uncontroversial that $a = b \Vdash \Box a = b$. The latter would lead one to expect that even if objects were to be viewed as functions from worlds to identities, then they would be constant functions, therefore, why think of them as functions to begin with? But recall that impossible and open worlds are not accessible by the relations on $\mathcal{W}$ that correspond to the modal operators; impossible and open worlds are accessible by intentional operators only. Hence, if we wish to maintain that $a = b \Vdash \Box a = b$, and that it is not possible for an object to have a different identity than it actually does, then it is easy to do so: keep the modal operators restrained to $\mathcal{P}$. However, according to Priest, it is very reasonable to suppose that at open worlds – which I emphasise correspond, so to speak, to intentional states – objects have different identities than they actually do. For example, the Hooded Man scenario is a case in point: for all you know, Bob and Clark are distinct individuals. That is, there is an open/ impossible world that is epistemically accessible to you where Bob and Clark are numerically distinct individuals, while it is nevertheless necessary that Bob and Clark are one and the same.

Formally these constraints are very easy to impose; we just want to make it the case that functions in D are constant over $\mathcal{C}$ but need not be constant over $\mathcal{O}$. We do so by requiring the following two conditions hold of our models:

- We want objects to have the same identity as in actuality at all closed worlds, hence:

$$\text{for any } d \in D \text{, if } w \in C \text{, then } d(w) = d(@).$$

- We also want functions to not discriminate between things that have the same identities at @, therefore we require the following:

  for any function f, and for any $d_i, e_i \in D$ for $i \in [1, n]$, if for all $i \in [1, n]$ we have that $d_i(@) = e_i(@)$, then $\delta(f)(d_1, ..., d_n) = \delta(f)(e_1, ..., e_n)$.

Priest 2005 proves in the Appendix to Chapter 2 that imposing the above two conditions on the semantics does not invalidate any instance of SI that is not an instance of SII – a proof which I do not repeat here. It remains to illustrate that the current framework is capable of invalidating unlovely instances of SII, such as the Hooded Man inference. This is simple. The inference in question is

$$b = c, t\Psi Hb \vdash t\Psi Hc.$$

To show that our semantics invalidates it we need only construct a counter-model such that $@ \Vdash^+ b = c, @ \Vdash^+ t\Psi Hb, @ \nVdash^+ t\Psi Hc$. The following is such a model:

$$\begin{aligned}
&\mathcal{C} = \{@\},\\
&\mathcal{O} = \{w\},\\
&Q = \{o_1, o_2, o_3\},\\
&D = \{d_1, d_2, d_3\},\\
&R_\Psi^{d_1} = \{(@, w)\},\\
&d_1 \text{ is such that } d_1(@) = d_2(w) = o_1,\\
&d_2 \text{ is such that } d_2(@) = d_2(w) = o_2,\\
&d_3 \text{ is such that } d_3(@) = o_2 \text{ and } d_3(w) = o_3,\\
&\delta^+(Hx, w) = \{o_2\},\\
&\delta^-(Hx, w) = \{o_3\}.
\end{aligned}$$

For consider that since $d_2(@) = d_3(@)$, it follows that $@ \Vdash^+ b = c$, and since $d_2(w) \in \delta^+(Hx, w)$ we have that $w \Vdash^+ H(b)$. But $d_3(w) \notin \delta^+(Hx, w)$, hence $w \nVdash^+ Hc$. Therefore, since $R_\Psi^{d_1} = \{(@, w)\}$, we have that $@ \Vdash^+ t\Psi Hb$ and $@ \nVdash^+ t\Psi Hc$ – as required.

With the semantics for a language with identity now fully in place, we now turn to a famous argument raised against theories of the extended possible worlds kind in Soames 1987, as promised in §2. After spelling out the argument I explain that it cuts no ice against Priest's framework developed above.

Soames first invites us to consider the following inference, call it the "Hesperus'/'Phosphorus' inference':

The ancients believed that 'Hesperus' referred to Hesperus and 'Phosphorus' referred to Phosphorus. (3.5)

The ancients believed that 'Hesperus' referred to Hesperus and 'Phosphorus' referred to Hesperus. (3.6)

The ancients believed that 'Hesperus' referred to Hesperus and 'Phosphorus' referred to Hesperus and, for some x, 'Hesperus' referred to x and 'Phosphorus' referred to x. (3.7)

The ancients believed that for some x, 'Hesperus' referred to x and 'Phosphorus' referred to x. (3.8)

According to Soames, any theory that satisfies the following assumptions must recognise the 'Hesperus'/'Phosphorus' inference as valid:

Relational Semantics for Ascriptions (RS) Propositional attitude ascriptions report relations to the semantic contents of their complements. That is, an individual 't Ψ's that S' is true relative to a context C, an assignment of value to variables s, and a world w if and only if in w the denotation of t under s R_{Ψ} to the semantic content of S relative to C and s, where if Ψ is the verb 'knows' then R_{Ψ} is the relation of knowing, if 'believes' of believing etc.

Direct Reference (DR) Names, indexicals, and variables are directly referential. That is, the semantic content of a name, indexical, or variable relative to a context C and an assignment of values to variables s, is the denotation of the term with respect to C and s.

Unstructured Content (UC) The semantic content of a sentence S relative to a context C is the set of worlds in which the S is true relative to C.

Distribution over Conjunction (DC) Many propositional attitude verbs, including 'say', 'assert', 'believe', 'know', and 'prove' distribute over conjunction. That is, let Ψ be any of these verbs, then if $t\Psi A \wedge B$ is true with respect to a context C, an assignment s and a world w, then $t\Psi A$ and $t\Psi B$ are true with respect to C, s and w.

SI in non-intentional contexts (SINI) If S and S' are nonintensional formulas with the same grammatical structure, which differ only in the substitution of constituents with the same semantic contents – relative to their respective contexts and assignments of values to variables–, then the semantic contents of S and S' will be the same, with respect to those contexts and assignments.

Truth Conditions for Conjunctions (TC) A conjunction $A \wedge B$ is true with respect to a context C, assignment s, and world w if and only if A and B are both true in w, relative to C and s. Thus, the semantic content of a conjunction, relative to C and s, is the intersection of the semantic contents of the conjuncts, relative to C and s.

Truth Conditions for Particular Generalisations (TP) A particular generalisation $\exists xFx$ is true with respect to a context C, assignment s, and world w if and only if there is some object o in w such that Fx is true at w with respect to C relative to an assignment s' that differs from s at most in assigning o as value of x. The semantic content of $\exists xFx$ relative to C and s is the set of worlds w such that for some object o in w, o satisfies Fx with respect to C, s, and w.

The argument is then that, given (3.5) is true and (3.8) is false, some principle out of the ones above list must be abandoned, given that the reasoning from (3.5) to (3.8) proceeds through RS, DR, UC, DC, SINI, TC, TP in valid steps – note that from (3.5) we move to (3.6) through UC, RS, DR, SINI and TC, from (3.6) to (3.7) through UC, DR, TC and TP, and from (3.7) to (3.8) through DC and TC. Soames then argues that the invalid step is that from (3.5) to (3.6), and that the culprit is UC hence taking the current argument to be a reductio of the thesis that propositional contents are to be identified with sets of worlds of any kind.

To be sure, Soames's argument is an argument against any theory that endorses RS to TP. And as we have seen Priest's semantics does not endorse most of those. As is clear from the truth-conditions for quantified formulas and conjunctions at open worlds, TC and TP do not hold. Open worlds also destroy SINI as all formulas are treated essentially as atomic there, and the same goes for DC since intentional operators are allowed to access open worlds in which conjuncts may be true but their conjunction false, and vice versa. Hence, given that one or more of these principles are essential for each and every deductive step in the inference from (3.5) to (3.8), Soames's argument does not raise any trouble against the current theory for each step of the 'Hesperus'/ 'Phosphorus' inference comes out as invalid. However, in Ripley 2012 the author argues that even if Priest's semantics were to be tweaked so as to satisfy TC, TP, SINI and DC – note they satisfy RS, DR, and DC already – contra Soames the 'Hesperus'/ 'Phosphorus' inference would still be blocked, and moreover, the inference would be blocked at the exact reasoning step Soames identifies as most controversial, the passing fro (3.5) to (3.6).

Here is then how Priest's framework regimented so as to validate all of Soames's assumptions blocks the inference from (3.5) to (3.6). For convenience, let h stand for 'Hesperus', h' for "Hesperus", p for 'Phosphorus', p' for "Phosphorus", R for the relation of 'refers to', t for an appropriate bunch of ancients, and let B be the belief operator. Then the claim is that if $@ \Vdash^{+} tB(h'Rh \wedge p'Rp)$ and $@ \Vdash^{+} h = p$, then $@ \Vdash tB(h'Rh \wedge p'Rh)$. Note that we do not simply want to present a counter-model. We want to present a counter-model that makes justice to the intended readings of (3.5) and (3.6). As Ripley puts it "[a]s Soames 2008 makes clear, Soames does not mean to claim that there are no circumstantialist countermodels to the inference from [(3.5)] to [(3.8)]. Rather, he means to claim that no circumstantialist theory (meeting his assumptions) can account for the truth of [(3.5)] and falsity of [(3.8)], *given what they actually mean*" (Ripley 2012, p. 115). Hence extra care is needed in constructing the counter model. To simplify the presentation we only consider two worlds – the actual and an open world – but for the rest the following counter-model serves our intuitions just right:

$$
\begin{aligned}
&\mathcal{C} = \{@\},\\
&\mathcal{O} = \{w\},\\
&Q = \{o_1, o_2, o_3, o_4, o_5\},\\
&D = \{d_t, d_h, d_p, d_{h'}, d_{p'}\} \text{ such that:}\\
&d_t(@) = d_t(w) = o_1,\\
&d_h(@) = d_h(w) = o_2,\\
&d_p(@) = o_2, d_p(w) = o_3,\\
&d_{h'}(@) = d_{h'}(w) = o_4,\\
&d_{p'}(@) = d_{p'}(w) = o_5.\\
&R^t_B = \{(@,@), (@,w)\},\\
&\delta^+(R,@) = \{(o_4,o_2), (o_5,o_2)\},\\
&\delta^+(R,w) = \{(o_4,o_2), (o_5,o_3)\}.
\end{aligned}
$$

Then, according to this model, since $d_h(@) = o_2 = d_p(@)$, we have that

$$@ \Vdash^+ h = p,$$

which serves right the intuition that Hesperus is Phosphorus since at all closed worlds – which happen to be just @ – Hesperus is the very same thing as Phosphorus. Moreover, since for x being either @ or w we have that $(d_{h'}(x), d_h(x)) \in \delta^+(R, x)$, it follows that

$$@ \Vdash^+ h'Rh \text{ and } w \Vdash^+ h'Rh,$$

and similarly since $(d_{p'}(x), d_p(x)) \in \delta^+(R, x)$, it follows that

$$@ \Vdash^+ p'Rp \text{ and } w \Vdash^+ p'Rp.$$

Now given we have assumed this regimented version of Priest's semantics satisfies TC, we conclude that:

$$@ \Vdash^+ h'Rh \wedge p'Rp \text{ and } w \Vdash^+ h'Rh \wedge p'Rp,$$

and since @ only accesses w and itself through R^t_B, we finally have that

$$@ \Vdash tB(h'Rh \wedge p'Rp).$$

However, $(d_{p'}(w), d_h(w)) = (o_5, o_2) \notin \delta^+(R, w)$ we have that

$$w \nVdash p'Rh,$$

and since $(@, w) \in R^t_B$ we may conclude that

$$@ \nVdash^+ tBp'Rh.$$

Hence, since we have assumed the semantics we are working with satisfy DC, we get that

$$@ \nVdash^+ tB(h'Rh \wedge p'Rh),$$

as required.

Let us take notice one last time that this counter-model expresses what we intuitively take it to be the case with the ancients: granted, Hesperus is Phosphorus, and granted the ancients have have the right beliefs about what 'Hesperus' and 'Phosphorus' refer to – they refer to Hesperus and Phosphorus respectively at each world they have epistemic access to. However, it is not believed by them that Hesperus is Phosphorus; that is, there is an open world they retain epistemic access to in which Hesperus is not Phosphorus. Hence, that very same world is a world in which 'Hesperus' and 'Phosphorus' are not co-refrential. Therefore, we have played Soames's game – have accepted his assumptions and have provided an intended model – yet we have shown that Priest's underlying framework is a circumstantialist framework – a framework abiding by UC – that nevertheless blocks the 'Hesperus'/ 'Phosphorus' inference just where Soames, and us intuitively, wanted it blocked.

Priest's semantics are a success in at least the following ways: the framework validates all instances of SI that are not instances of SII, while instances of SII – such as the Hooded Man Paradox – can be blocked, and further it is immune to Soames's argument. However, I wish to argue, the success is mainly owed to a technically sound feature of the semantics that is nonetheless philosophically dubious. As we have seen, the interpretation of $\mathcal{L}^+$ is a septuple which includes two domains, one of identities Q, and one of objects D, where objects are viewed as functions from worlds to identities. And as we have established, the addition of Q and re-interpretation of D as functions from W to Q suffice on a technical level to serve Priest's aims. My initial concern is, as Hale concisely puts it, that "it is far from clear what the elements of Q – 'identities' – are supposed to be" (Hale 2007, p.103), which in turns renders it unclear what the elements of D – 'objects' – really are. Priest's system takes its cue from contingent identity systems, such as the one developed in §18 of Hughes and Cresswell 1996. Although the systems are not identical in all their details, they have enough in common to draw a comparison. The authors there motivate their treatment of objects as functions as follows: think of a name or definite description as "standing for a single 'object', though one which in a more usual sense of 'object' may be one object in a certain situation but a different one in another", where "[s]uch 'objects' are often called intensional objects or individual concepts" (Hughes and Cresswell 1996, p.332-3). This exposition already makes some philosophical sense, for talk of individual concept is rather reminiscent of Fregean senses. However, as Priest notes in a footnote:

> It is tempting to think of identities as Fregean senses, but this would not be right. If anything, it is the members of D that are more like senses, since they determine behaviour across worlds. [...] It would also be a mistake to interpret members of D, in the present semantics, as senses, however. They are simply the objects themselves. (Priest 2005, p.44)

The complaint here is, simply, that Priest offers little to no explanation of what 'identities' and 'objects' are meant to be.

But moreover, the more one looks into their behaviour, the more puzzling things are. First of all, note that the denotations of names are members of D i.e. objects, and therefore, you, me, and whoever/ whatever else we can name are all distinct from our identities. On the most natural reading of 'identity' – and given Priest has not given much else to go on – this is simply puzzling. A puzzlement that carries over to Priest's 'objects', for we have been told that objects are functions from worlds to identities, and just what an identity is remains unclear. If Priest wants to offer more than a technical fix to issues arising from the apparent failure of SI in intentional contexts, then a philosophically illuminating or at least convincing account of Q is in order.

Other than this initial puzzlement over the nature of the members of Q, Hale 2007 raises a point that casts further doubt over the philosophical substance of Priest's solution to the aforementioned problems. Priest's solution to issues arising from SII rely critically on the fact that there is no way to refer to elements of Q; if there was a way to do so, and assuming there could be more than one name denoting a single element of Q, we could get revenge problems. But we may drive the point slightly differently, for Priest's solution works by postulating a class of entities that speakers are completely unaware of and incapable of denoting, and re-interpreting ordinary names as designating functions from worlds to such entities. Granted, the solution works on a technical level, but it certainly does not cast much light on the fact that the ancients knew Hesperus was self-identical but did not know that Hesperus was Phosphorus. All it tells us is that the names 'Hesperus' and 'Phosphorus' denote functions such that there is a world epistemically accessible to the ancients where the image of one function is not the image of another, where the 'identity' of Hesperus is not the 'identity' of Phosphorus, but what an 'identity' is neither we nor the ancients know.

Lastly, the following is an odd feature of the account: the extensions of names are elements of D, while the extensions and anti-extensions of n-place predicates are subsets of Q^n. That is, names denote 'objects', while predicates hold of 'identities'. Sure enough, the technical details pan out for when it comes to evaluating whether an 'object' lies in the extension of a predicate it is always relative to a world, and when you combine 'object' and world you get the image of the function that is the object, that is, its 'identity'. However, we are trying to capture meaning and intentionality, and it is very odd to think that when I say 'Amy is dancing' what I am doing is predicating a property of an entity I have no means of picking out – i.e. the 'identity' of Amy – while using a name of an entity that is not of the sort which falls under extensions of predicates. All in all, none of these are damning criticisms, and perhaps the view could be elaborated so as to be more philosophically substantial, but as things stand it is not. What I have set out to find in the current thesis is a philosophically substantial and plausible account of meaning, and not merely a model that works.

In the current section I built towards Priest's extended possible world semantics for a first order modal language with identity and intentional operators. We saw that the framework destroys unwanted principles such as logical omniscience and closure under entailment, as well as accommodates for our Fregean intuitions in cases where identity interacts with intentionality. Lastly, we saw that not only are the semantics not touched by Soames's influential 1987 argument, but that the theory developed can be used to construct a counter-example to Soames's overarching argument. However, there were two concerns raised: firstly, that the semantics's success with Soame's argument and SII seems to rely on a technical 'trick' that is philosophically rather puzzling, and secondly, that it seems too cheap – it destroys too many inferences concerning intentional operators. We now turn to Jago 2014 who is motivated by this latter concern and constructs an alternative extended possible worlds theory that attempts to strike a balance between validating too many or too few inferences concerning intentional operators, all the while avoiding trouble from SII in intentional contexts without appealing to a questionable, Priest-like technical artefact.

3.3 Jago

3.3.1 The Problem of Bounded Rationality

In the current section we look closer at the tension previously only hinted at. Jago dubs this 'the problem of bounded rationality', and describes it as "the problem of accounting for the normative connections between logically-related contents whilet avoiding logical idealisation" (Jago 2014, p. 163). After shorty acknowledging that Priest's semantics is an unsatisfying solution to the problem, we turn to examine Jago's own response.

As we have seen, proponents of structured propositions and Priest alike are very preoccupied with intentionality in their attempts to construct theories of content. That is, when aiming for a theory of content, a central and often tough challenge to be met is giving an account of epistemic and other intentional contents. As we have discussed time and again, standard possible world approaches to modelling epistemic agents render such agents *ideal*, insofar as the frameworks predict that agents are logically omniscient – believe and know all logical truths – and that their sets of beliefs and knowledge are closed under logical entailment – agents believe/ know all logical consequences of what they believe or know. Real epistemic agents are simply not like that. So one desideratum is to not give a completely idealised account of human rationality.

However, I have spoken of tension. The tension becomes clear as soon as we acknowledge that we do not want to give an overly illogical account of human rationality either. As Jago invites us to do, let us ponder for a second on how odd it would be to ascribe to an agent a belief in $A \wedge B$ but not a belief in A, a belief in $A \rightarrow B$ and A but not in B, or a belief in A but not in $A \vee B$. Of course I am not claiming that it is impossible for agents to hold a belief in, say, $A \wedge B$ but not in A, or that if they believe that $A \wedge B$ then they *must* believe that A; the claim is rather that ascribing an agent such a belief system – saying that agent a believes that $A \wedge B$ but does not believe that A – seems very strange and highly counterintuitive. The reason it strikes us as such, Jago suggests, becomes clearer if we endorse a behaviour-based approach to epistemology, that is if we consider that "a belief attribution is part of a holistic attempt to make best sense of the agent's behaviour" (Jago 2014, p. 164) – where same goes for knowledge. We may then say that when attempting to make best sense of an agent that believes that $A \wedge B$, it is most intuitive that this will involve us in ascribing her a belief that A, and ascribing her a belief that B. To take an example, if Batman announces 'the Joker is a villain and Catwoman is a villain', the people of Gotham would expect him to take one on and then the other, a behaviour best made sense of by ascribing two beliefs of him: that the Joker is a villain, and that Catwoman is a villain.

But there is some idealisation here, for sometimes epistemic agents may seem to believe or may claim to believe that $A \wedge B$ but not that A, that $A \rightarrow B$ and that A but not that B, that A but not that $A \vee B$. What should these empirical data mean for us? Jago argues that in such cases it is most tempting to understand the agents in question as confused about the meaning of the connectives, or as in some other sense mistaken – just as in a case of being given wrong change we do not assume that the well-meaning cashier is confused about the meaning of basic arithmetic symbols or numbers, but rather that they performed some mistake in calculation. Hence the suggestion is that we acknowledge that competent language users make mistakes even when employing such simple words as the connectives. However, this should not matter much for our theory of meaning; when we catch speakers red-handed, we need not take their statements at face-value, apply a disputation principle and attribute to them a highly unlikely epistemic status. Instead, on the assumption that a belief ascription is part of a holistic effort to make best sense of an agent's behaviour, and given that there might be cases that cannot be made sense of under this type of explanation – as for example cases in which agents seem to be misusing the connectives – we must afford some idealisation. This is an idealisation away from an agent's actual behaviour or assertions, and, focusing on what matters for current purposes – the logical connectives – this means that from the outset we recognise there are strong normative links between $A \wedge B$ and A, A and $A \vee B$, and $A \rightarrow B$, A and B – odd cases notwithstanding. Therefore, our second desiredatum is that some simple inferences hold true under the scope of epistemic operators, as in – to borrow Priest's language – $t\Psi A \wedge B \Vdash t\Psi A, t\Psi A \rightarrow B, t\Psi A \Vdash t\Psi B, t\Psi A \Vdash t\Psi A \vee B$.

But now the tension between the deseridata should be obvious: we do not want logical omniscience and closure under entailment, but we do not want wildly illogical agents either; we want to strike a balance; we want a suitable degree of idealisation. As Jago puts it, the problem of bounded rationality is "the conflict between normative principles of rationality and the fact that the agents with which we are concerned have limited cognitive resources" (Jago 2014, p. 165). And phrased as such, it is easy to see that the problem of bounded rationality is not limited to modelling human intentionality. The issue reproduces on the level of cognitive significance of sentences or utterances thereof, and the latter being equally central to an account of content as epistemic content is, we shall be much concerned with it too.

To see how the problem of bounded rationality creeps up on the level of content, consider Frege's puzzlement with informative identity statements. Let us here use the 'surprise test', as in, let us assume than if a sentence or utterance thereof is surprising, then it must be informative. That Hesperus is Phosphorus surely surprised ancient astronomers as it overturned a long-standing tradition that would have two astronomical bodies in the place of one. However, according to standard possible world frameworks, 'Hesperus is Phosphorus' being a necessary truth it is not informative. Or according to any theory that insists 'Hesperus is Phosphorus' and

'Hesperus is Hesperus' express the same proposition – and given the reasonable assumption that a thing being self-identical comes as no surprise – then, again, 'Hesperus is Phosphorus' comes out uninformative. Or, assuming mathematical truths are necessary, the standard possible worlds framework would have it that no logical or mathematical theorems are informative. But this is certainly not true: there are surprises in math and logic, and to insist that all logical and mathematical endeavours can at best lead to something uninformative – so long as mathematicians and logicians are striving to establish logical and mathematical truths – would leave a lot of people without a job. So, on the one hand, we want an account of content that allows some necessary truths (and falsehoods) to be cognitively significant.

But on the other hand, we would not want all necessary truths and falsehoods to count as informative. Some, such as arguably $A \vee\ /A$ or $1 + 1 = 2$ seem trivial, almost definitional of the symbols embedded in them. And relatedly, given that logical truths or theorems can be seen as results of zero-premise deductions, we may want to look into what counts as an informative step in a deduction. Surely the step from $A \wedge B$ to A, from $A \rightarrow B$ and A to B, or from A to $A \vee B$ are as trivial as they get. But this is puzzling for proofs leading to the establishment of surprising logical results proceed through a series such trivial steps, and it is hence hard to see how to give an incremental account of their cognitive significance. We thus have another instance of the problem of bounded rationality: normative principles connecting, for exampe, $A \wedge B$ to A, $A \rightarrow B$ and A to B appear to dictate that such inferences are trivial and uninformative, yet empirically for agents with limited cognitive capacities such as ourselves, some logical truths and some valid deductions are informative.

Priest's semantics are not a solution to the the problem of bounded rationality. Granted, they do satisfy the first desiredatum that epistemic agents are not logically omniscient, that their beliefs are not closed under entailments, that logical truths and valid deductions can be informative – so long as we identify a proposition with the set of worlds which make it true, both open and closed. However, as the account stands it does not satisfy the second desiredatum: we have no guarantee that agents know even the simplest logical truths, that their beliefs are closed under any kind of entailment, that some logical truths and valid inferences are by every means trivial. This is precisely because Priest invokes *open* worlds – worlds that are completely anarchic, not closed under any rue of inference except identity and thus making true any set of formulas whatsoever – and allows such worlds to be epistemically accessible by agents. We now turn to Jago's own account which supposedly dispenses with the problem of bounded rationality.[8]

Although Priest's semantics are allegedly compatible with a variety of metaphysical stances towards possible and impossible worlds, Jago's is not. A short discussion of the metaphysics of impossible worlds in general and according to Jago is to be found in §3.4 since I wish to keep such issues out of the main focus. For now, let us assume that Jago has somehow constructed a totality of worlds that suit his purposes and let us turn to looking into his account of the epistemic space, and subsequently examine the formal framework he develops to model epistemic content and which purportedly dispenses with the problem of bounded rationality.

3.3.2 The Epistemic Space

We have seen how Jago constructs and metaphysically motivates his totality of worlds – possible and impossible. Jago wants to use this totality of worlds to built a semantics for epistemic content but for this aim the unstructured set of worlds will not do. As we have seen, Priest has

[8] I gave it a lot of thought and decided to not go through the 'selected worlds' strategy so as to transition between bounded rationality / Priest / Jago. If you think it would be better to do so please let me know!

a totality of worlds that includes many impossible ones. And Jago does too, but Jago wants to address the problem of bounded rationality. And it is his main thesis that imposing structure on the totality of worlds – resulting in what he calls the epistemic space – is key in doing so. Jago does not want to impose logical structure on each world in the totality of worlds and then select a subset of the worlds to render epistemically accessible. Instead, what he wants is to impose logical structure on the totality of worlds. This is then what the current subsection is devoted to.

Recall the problem of bounded rationality: to destroy epistemic idealisation we must invoke logically anarchic worlds, but to avoid trivialising the meaning of the logical constant we must deny that all such worlds are epistemically accessible. The obvious move here is to select some but not all worlds to be epistemically accessible. However, Jago does not wish to single out the epistemically possible worlds by imposing on them logical constraints, as for example by demanding that any world that is epistemically accessible represents that A, if it represents that $A \wedge B$. As Jago discusses in many points across chapters 6 and 7 of Jago 2014, such an approach leads to insurmountable difficulties. Jago's strategy is instead to impose structure on the totality of worlds, and select epistemically possible worlds based on their position in the structure.

More precisely, Jago argues that the totality of worlds is governed by normative principles – although individual worlds are not – and that the set together with this structure comprises the epistemic space. These normative principles are principles that preserve the meanings of the logical connectives by 'representing' inference rules, as follows. Let us take the simple rule of conjunction introduction, $A, B \vdash A \wedge B$. Now let us take a world w such that w represents that A, and it represents that B but it does not represent that $A \wedge B$ – note that we have imposed no constrains on what sets of the world-making language constitute a world, and hence such a world as w is to be found in our collection. Then, a world w' just like w which differs only insofar as it represents additionally that $A \wedge B$, is a world that w is related to, where the relation in question is a irreflexive, asymmetric and intransitive relation. Do that for every world and any A, B and we have a relation on our totality of worlds that may be seen as capturing "the rational relations between representations established by *conjunction introduction*" (Jago 2014, p.198). Note that it is not the worlds *per se* that are closed under conjunction introduction, instead the normative force of the rule is captured by the relations between our worlds/ representations.

Now, each arrow represents a one-step-inference to a single conclusion of the sort that we often consider trivial and uninformative. However, assuming we have arrows for all inference rules of an appropriate proof system, chains of worlds represent longer proofs of the sort we often consider illuminating.[9] This is anticipating a bit, but given the epistemic space - the totality of worlds and the structured imposed on them the way described for conjunction introduction, but now expanded to include all other inference rules – we may expect to solve the problem of bounded rationality for we have essentially an order on worlds. This order can help us explain why some inferences are informative while others not – loosely speaking the former unwind through longer paths across the epistemic space – why some necessary truths are informative while others not – again, informative necessary truths take longer to establish – or why some inconsistencies are epistemically accessible while others are too blatant to be so – for example, worlds representing blatant contradictions, or contradictory worlds whose inconsistencies can

[9]It is a big issue what proof system to pick, and many an inference rules – such as rules involving assumption-making or reasoning by cases – are harder to translate into relations between worlds in this manner. Moreover, there are deductively equivalent proof systems that will result in a different epistemic space, which might lead one to think that a choice between such systems renders our structure arbitrary in a way it should not be. All these issues I have decided to completely black-box. The interested reader should consult §7.4 of Jago 2014. This is not a place to delve into the details, important as they may be, but I believe the idea behind Jago's framework is intuitive enough to follow in their absence. (If Franz, Luca you disagree then I can go into it.)

be uncovered in just a few steps are themselves fairly blatant too and hence not epistemically accessible.

As the reader may already have suspected, according to Jago it is not always a determinate matter whether a given world is epistemically possible or not; we have a variant of the sorties paradox. A world that represents a blatant contradiction is not epistemically accessible. A world whose contradiction is uncovered in one step through epistemic space is not epistemically accessible. A world whose contradiction is uncovered in one more step than that of a world that is not epistemically accessible is not epistemically accessible. But there are epistemically accessible inconsistent worlds. Jago insists that your favourite account of vagueness will work for the current case as well as for any, and that it not being a fully determinate matter for all worlds whether they are epistemically accessible fits with the empirical facts. So Jago is content with the following picture: a that represents that A and that represents that $\neg A$ for some A, is a blatantly inconsistent world, and such worlds are not epistemically accessible. A world that belongs to a path leading to a blatantly contradictory world is an inconsistent world that might or might not be epistemically accessible. If it is far away enough, it is a world that represents only a subtle contradiction and is thus epistemically accessible. If it is very close, it is an epistemically inaccessible world. If it falls somewhere in the middle – and here enters talk of vagueness – it is a borderline epistemic possibility. In the following section we present the framework based on Jago's epistemic space more precisely, and look at how it models informational content and epistemic content, focusing on how it addresses the issue of bounded rationality.

3.3.3 Modelling Content

Jago has constructed a highly structured epistemic space, structured under proof rules capturing normative principles pertaining to the meanings of logical connectives. The next step is to use this epistemic space to construct notions of content. The contents we will be focusing on are those in play when we discuss knowledge, belief, information and cognitive significance; we refer to all these under the umbrella term 'epistemic contents'.

The simplest notion of content – taking its cue from standard possible world semantics – is that which we get by identifying the content of 'A' relative to some context of utterance with the set of all epistemically possible worlds representing that A. This is already a more fine-grained notion of content than that the SPW delivers, for epistemically possible worlds are more fine-grained than logically and metaphysically possible ones. Moreover, we may consider to not identify contents with sets *simpliciter*, but with sets plus the structure they inherit from the epistemic space – that is, we identify contents with regions of the space which are themselves normatively structured. Moreover, given we have acknowledged that it is not always a determinate matter whether a given world is epistemically possible or not, and given that w belongs to the content of 'A' if and only if (i) w is epistemically possible, and (ii) w represents that A, we must conclude that contents also inherit the 'vagueness' of the epistemic space.

We may enrich this simple account of content many ways, but one that will be particularly useful is for our account of content to treat both truth and falsity on par. That is, the content of 'A' relative to a context of utterance will be a tuple $\langle [[A]]^+, [[A]]^- \rangle$, where $[[A]]^+$ is the positive and $[[A]]^-$ is the negative content of 'A'. The positive content is as above, the region of the epistemic space consisting of the epistemically possible worlds that represent that A. The negative content is similarly the region of the epistemic space consisting of the epistemically possible worlds that do not represent that A. That is, when 'A' is a sentence of out world-making language, and w an epistemically possible world, we have that $w \in [[A]]^+ \iff 'A' \in w$ and $win[[A]]^- \iff 'A' \notin w$. But on top of defining contents as pairs of sets of epistemically possible worlds, we may define more complex structures on our epistemic space to be used for giving an account of the epistemic and doxastic states of agents, much like it is done in epistemic logic using accessibility relations. However, before doing so we briefly look into valid inferences and how we may distinguish between the informative and uninformative ones, for this distinction becomes useful when giving an account of epistemic-doxastic states.

The current aim is to account for the idea that, intuitively, some valid inferences are informative while others not, in line with empirical data of the sort that a deductive move from $A \wedge B$ to B strikes one as rather trivial, while the proof of Fermat's Last Theorem does not. However, every single step is the latter proof is as trivial as the deductive step that is the former proof. Hence, a purely incremental account of informativeness will not do, neither will it do to pick one step in the deduction and locate the proof's informativeness there, for the question arises anew, what is it that makes that step informative? Jago, in seeking a solution from his epistemic space, first tries out a notion of informativeness familiar from SPW: informativeness is a matter of ruling out epistemic possibilities, that is, A is informative if upon coming to know that A the set of epistemically possible worlds shrinks, or else, the content of a valid deduction $A \vdash B$ is the difference between the set of epistemically possible worlds assuming the agent knows A but not B, minus the set of epistemically possible worlds assuming the agent knows in addition that B, so that a deduction is informative if and only if its content is non-empty. The issue of adhering to such a naive notion of informativeness is that, given many an incomplete worlds are deemed epistemically possible, far too many deductions are rendered informative such as all those from $A \wedge B$ to A.

The solution to this is, according to Jago, to analyse the content of a valid deduction $A \vdash B$ as $[[A]]^+ \cap [[B]]^-$, where if A is a set of premises $\{A_i, ..., A_n\}$ $[[A]]^+ = \bigcap_{i \in [1,n]} [[A_i]]^+$, and if B is a set of conclusions $\{B_1, ..., B_n\}$, we have that $[[B]]^- = \bigcup_{i \in [1,n]} [[B_i]]^-$. Hence, the proposal goes, what we are checking in order to see whether a valid deduction is informative is whether the truth of the premises is trivially incompatible with the falsity of the conclusions. If they are trivially incompatible – that is if no epistemically possible world lies in $[[A]]^+ \cap [[B]]^-$ – then the inference is uninformative. If the content is non-empty, then the inference is informative. Jago's diagnosis is that the naive notion of informativeness of a deduction $A \vdash B$ failed, for the worlds in $[[A]]^+ \setminus [[B]]^+$ do not take a stand on whether B. However, the current analysis of content which would have the content of the deduction be $[[A]]^+ \cap [[B]]^-$, does take a stand on whether B, therefore if $[[A]]^+ \cap [[B]]^-$ is empty, then any world according to which the premises are true while the conclusions false is not an epistemically possible world, it is a blatantly inconsistent one, which in turn implies that there is no information to be gained passing from the premises to the conclusion. This is then the account of informativeness that Jago endorses: the content of a valid deduction is modelled as the combined truth of the premises and falsity of the conclusions – contra SPW – yet the SPW idea that a deduction is informative only if there is informational gain to be had from passing from its premises to its conclusion is retained.

This bring us closer to what we need in order to model epistemic and doxastic states of agents. However, there is one last worry we need to address. This worry rises from precisely the fact that Jago does not want to close worlds under inference rules that 'define' the meaning of the connectives, but instead tries to capture these normative principles by imposing structure on the whole space. More specifically, in a SPW setting, the meaning of the logical connectives are captures by means of content-inclusion relations. That is, focusing on conjunction we have that $[[A \wedge B]]^+]] \subseteq [[A]]^+$, and on implication $[[A \rightarrow B]]^+ \cap [[A]]^+]] \subseteq [[B]]^+$. This is not something we have in the current setting where plenty of open worlds in Priest's sense are deemed epistemically possible. However, albeit not as straightforwardly as appealing to content-inclusion as above, we do have a way to capture on the level of content and through inclusion these normative principles that were the second desideratum contributing to the tension the problem of bounded rationality exemplifies. This is how: let W_e be the set of all epistemically possible worlds, and define $[[A]]^{-c} := W_e \setminus [[A]]^-$ such that $[[A]]^{-c}$ is just the set of all epistemically possible worlds that do not represent that A is not the case. Thus, if we were limited to possible, closed worlds, $[[A]]^{-c}$ would just be $[[A]]^+$, but we are not – we do allow open worlds. Hence, the above content-inclusion principles simply take the following form:

$$[[A \wedge B]]^+ \subseteq [[A]]^{-c} \tag{3.9}$$

$$[[A \rightarrow B]]^+ \cap [[A]]^+ \subseteq [[B]]^{-c} \tag{3.10}$$

To see how the current framework validates (3.9), suppose for *reductio* that $[[A \wedge B]]^+ \nsubseteq [[A]]^{-c}$. This means that we have some epistemically possible world w such that $w \in [[A \wedge B]]^+$ but $w \notin [[A]]^{-c}$, i.e. $w \in [[A]]^-$. But then there is a one step path through the epistemic space, that there is a world w' that w is directly connected to, such that $w' \in [[A]]^+$ – given that $A \wedge B \vdash A$ is an inference rule in any appropriate proof system. Clearly w' is an epistemically impossible world – $w' in [[A]]^+$ and $w' \in [[A]]^-$ – and no such world can be reached by a single step from an epistemically possible world. Therefore, w is epistemically impossible: a contradiction.

The problem of bounded rationality has been addressed, as far as content is concerned. We now turn to Jago's account of epistemic and doxastic states, and how it deals with the second facet of bounded rationality. I do not go as far as giving the full semantics here, but do sketch how it may be developed so as to account for all worries Jago voices; to do otherwise would simply take us too far afield.

Jago's semantics for knowledge and belief takes its cue from standard SPW epistemic logics. However, from the very outset, Jago is playing on a much more fine-grained and structured field, that of his epistemic space instead of the totality of possible worlds. And it is mainly this that allows for his account of epistemic-doxastic states to overcome issues of bounded rationality facing accounts of the SPW sort. More precisely, in a familiar move, given an agent a, we model her epistemic state using an epistemic accessibility relation R_a^K on the space of epistemically possible worlds, for sort W_e. A logically possible world w represents that agent a knows that A if and only if $(\forall w' \in W_e)((w, w') \in R_a^K \rightarrow \text{'}A\text{'} \in w')$, that is, if and only if all epistemically accessible, epistemically possible worlds represent that A. Epistemic states of different agents are represented using different accessibility relations on W_e, and given knowledge is traditionally viewed as factive, we impose the condition that accessibility relations must be reflexive. For the doxastic state of an agent a, we use a doxastic accessibility relation on W_e, R_a^B, which, given belief is non-factive, is non-reflexive. A logically possible world w represents that an agent believes that A if and only if all doxastically accessible epistemically possible worlds represent that A, or formally if and only if $(\forall w' \in W_e)((w, w') \in R_a^B \rightarrow \text{'}A\text{'} \in w')$. We say that an agent knows that A or believes that A if and only if all logically possible worlds represent that the agent knows or believes that A respectively.

Logically impossible worlds represent agents' beliefs and knowledge in an arbitrary manner, much like logically impossible worlds represent that A in arbitrary manners for logically complex sentences 'A' of other sorts – as for example worlds that represent that $A \wedge B$ but not that A. This anarchic behaviour is desirable so as to accommodate for facts of the following sort. We have assumed that knowledge is factive, so it is not possible for an agent to know something that is false. But that just means that there is an impossible world which represents that an agent knows a falsehood. Or, given our definition of knowledge above, if an agent knows that A, then it logically follows A is epistemically accessible for the agent, and thus epistemically possible. But then there must be a logically impossible world that represents an agent knowing or believing some A, for an A that is epistemically impossible. A world, considered as a pair $\langle w^+, w^- \rangle$, is logically possible if and only if, in addition to constrains of closure under entailment, consistency and completeness, and writing '$K_a A$' for 'agent a knows that A' it satisfies the following four constrains: (i) '$K_a A$' $\in w^+ \iff (\forall w' \in W_e)((w, w') \in R_a^K \rightarrow$ 'A' $\in w'^+)$, and (ii) '$K_a A$' $\in w^- \iff (\exists w' \in W_e)((w, w') \in R_a^K \wedge$ 'A' $\in w'^-)$, and similarly two constraints pertaining to belief using R_a^B instead. Worlds that fail to satisfy these constraints are impossible, and may represent both that $K_a A$ and that 'A' is false, for example.

With these minimal comments on the model of epistemic-doxastic states that Jago introduces, it is already clear how logical omniscience and closure under entailment are destroyed: as was the case for Priest, Jago allows for open worlds and for impossible worlds that make necessary truths false, and allows that some such worlds are epistemically possible and accessible. But Jago has set another goal for himself, namely to account for the fact that agents must know and believe certain truths merely in virtue of being rational, as is the case for knowing trivial consequencees of what they know or believing obvious truths. However, as noted when discussing the problem of bounded rationality, rational agents often enough fail to believe/ know obvious truths or truths that trivially follow from what they believe or know - recall we interpret a competent speaker who asserts an obvious falsehood as being temporarily confused about the meanings of the logical connectives, much as we interpret the well-meaning cashier who gives wrong change not as completely ignorant of basic arithmetic but as mistaken. Such instances Jago calls epistemic oversights and wants his framework to be able to account for. This is because the following is often the case: the knowledge of agents is not deductively closed although many of deductive rules make for trivial inferences. However, many a non-trivial results can be established through inferences all of whose steps are trivial. Yet we want to allow that just because such inferences may be long or may not concern the agent enough for her to go through with them, she may know the premises but not the conclusion. In such cases we may invoke epistemic oversight to the rescue, so that such an agent is not considered irrational. To account for such cases Jago appeals to the fact that his epistemic space incorporates some sort of vagueness, that is it is not always a determinate matter whether a given world is epistemically possible or not. This vagueness is inherited by epistemic accessibility relations – for epistemically accessible worlds are epistemically possible worlds that satisfy some further condition – and it is this indeterminacy in epistemic accessibility that he exploits to account for the fact that it is highly counter intuitive to ascribe certain epistemic oversights to agents – such as that they know $A \wedge B$ but not A – and not others. This is something certainly worth noting, but the line is drawn here, before the full-blown model that accounts for these is developed.[10]

Although not fully developed, in this section I have sketched jago's extended possible world semantics for modelling content and epistemic-doxastic states of agents. The motivating concerns behind Jago's highly sophisticated account are captured well by Lewis:

> I agree with [Priest] about the many uses to which we could put make-believedly impossibilities, if we are willing to use them. The trouble is that all these uses seem to require a distinction between the subtle ones and the blatant ones (very likely context-dependent, very likely a matter of degree) and that's just what I don't understand. (Lewis 2004, p.177)

The many uses Lewis is hinting at here, and why they require such distinctions, have been explained in our discussion of the problem of bounded rationality: epistemic content, informativeness, cognitive significance, epistemic-doxastic states all seem to require a theory that strikes

[10]The interested reader should look at §8.4 – 8.5 of Jago 2014.

a balance between ideal and completely anarchic rationality. Jago is attempting just what it is that Lewis claims to not understand. The most central move Jago makes is structuring the totality of worlds according to normative principles inherited on the structure by means of inference rules of an appropriate proof system, and not closing worlds under entailment so as to avoid idealisations. With this structured epistemic space in place, Jago can model epistemic content and epistemic-doxastic states in the same spirit – although not in perfect analogy – as strategies employed by SPW frameworks, and still manage to overcome issues pertaining to bounded rationality that doom the SPW approach.

3.4 Conclusion

In this section I summarise how Priest's and Jago's strategies fare with the data from §1, to see whether the extended possible worlds strategy succeeds in delivering the finess of grain a good theory of meaning should be expected to. Next, I briefly address the following: impossible worlds seem to raise controversial issues in metaphysics, and although this often weighs against them, I do not take it to be a major concern given the focus of the current thesis. Weighing in all the evidence, I conclude that extended possible worlds semantics, and in particular Jago's highly sophisticated account of epistemic content and epistemic doxastic states, is a very promising strategy altogether.

3.4.1 The Data and the Verdict

Many of the data have been discussed throughout the exposition of Priest's and Jago's accounts, so the current will serve as a summary of what is spread throughout the chapter. It serves well to organise things this way and have a comprehensive outlook of the theory's ability to draw hyperintensional distinctions, but a few additional data will be considered, namely indicative, subjective and counter possible-conditionals.

To begin with, both Priest and Jago allow for sufficiently many open worlds so that distinct necessary truths are distinguished from one another. Take for example (1.1) and (1.2). Given that a proposition is identified with the set of worlds that make it/ represent it as true, and given we allow for mathematically impossible worlds, there is a world w such that $w \Vdash 2+2=4$ and $w \vdash 3+3 \neq 6$, and such a world belongs to the proposition expressed by the former but not by the latter sentence. Hence, since propositions are sets and we have identified a world that is an element of one but not the other, it is clear that these propositions are not identical. According to Priest, the very same goes for logical impossibilities, since he allows for completely anarchic open worlds. However, according to Jago not all Priest-open worlds are epistemically accessible worlds – for example worlds where blatant contradictions are true – and propositions are thus identified with sets of epistemically accessible worlds. Therefore, according to Jago, some logical truths such as the ones expressed by $\neg(A \wedge \neg A)$ and $\neg(B \wedge \neg B)$ express the same proposition, precisely because open worlds where one or the other fails are not epistemically accessible; that is, we do not have a world where one is true while the other not so as to distinguish them set-theoretically for we have no world in which either fails makes it into the epistemic space. Arguably, and Jago has most certainly argued for that, this is no bad thing. These are trivial truths and as such can be seen as being epistemically on par, and the cost of distinguishing them by allowing all the open worlds is too great: we face the problem of bounded rationality.

Truths of the sort 'Hesperus is Phosphorus' and 'Hesperus is Hesperus' are easily distinguished too, for there are impossible worlds – that are epistemically accessible for Jago – in which Hesperus is not Phosphorus, while Hesperus is still self identical. Such worlds belong to the proposition expressed by the latter identity statement but not to those expressed by the former, and therefore, simply because of having different elements the two propositions are distinguished. Most necessary falsehoods are also distinguishable by the very same set theoretic means. Take (1.9) and (1.10) for example; there are possible worlds where Amy is the only person who managed to square the circle, and these are worlds that belong to the proposition expressed by the former but not by the latter sentence. Mathematical falsehoods and many a logical falsehoods are distinguished in a similar vein. However, we need note again that where contradictions of the sort $P \wedge \neg P$ are concerned, Priest's and Jago's opinions diverge. Priest allows for worlds instantiating blatant contradictions, and different worlds instantiate different blatant contradictions, hence no blatant contradictions are conflated. However, Jago does not allow worlds representing blatant contradictions into his theory of epistemic content, for he does not consider such worlds to be epistemically possible. Therefore, just as was the case for obvious/ trivial truths, these obvious, blatant falsehoods are conflated and viewed as expressing the very same proposition. As noted above, there is a tradeoff here: we either follow Priest and are weighed down but the problem of bounded rationality, or we follow Jago and equate all blatant contradictions. Although the scale seems to be arguing for Jago, it need be noted that there is something unsatisfying about holding that 'This ball is and is not round' and that 'In this picture it is both snowing and not snowing' express the same proposition. What I am after is an account of content *simpliciter*, and even if we grand Jago that all blatant contradictions and all trivial truths are epistemically on par, the point remains that the sentences expressing them do not strike us as saying the same thing – one is about sporting equipment, the other about the weather.

Lastly, for the sake of completeness, let us note one more time that the models for epistemic-doxastic states based on an impossible worlds framework do truth conditionally distinguish between propositional attribute reports embedding necessarily equivalent clause. This is owed to the fact already noted that the frameworks already distinguish between necessarily equivalent propositions, for such propositions are identified with sets of worlds that are much more fine grained than those employed in SPW semantics. To take an example, we have already distinguished between the propositions expressed by 'Hesperus is Phosphorus' and 'Hesperus is Hesperus', and given that 'The ancients believe that Hesperus is Phosphorus' and 'The ancient believe that Hesperus is Hesperus' are true if and only if all worlds epistemically accessible by the ancients belong to the proposition expressed by 'Hesperus is Phosphorus' and by 'Hesperus is Hesperus', respectively, it is clear that the former doxastic state is not equivalent to the latter. As far as accounting for epistemic-doxastic states go, Priest is already an improvement on the SPW approach for his agents are not logically omniscient and their knowledge and beliefs not closed under entailment. Jago does even better for his agents are not completely irrational either: they do not believe in blatant contradictions and most of the times believe and know trivial consequences of what they believe or know.

Recall that a datum the proponents of structured propositions could not accommodate for is the intuition that the following two indicative conditionals express distinct propositions, the former striking us as true while the latter as false: (1.19) 'If Hesperus not Phosphorus, then Hesperus is not Phosphorus', (1.20) 'If Hesperus is not Phosphorus, then Hesperus is not Hesperus'. Both Priest and Jago do justice to the above intuitions. (1.19) comes out true since even Priest's open worlds are closed under identity – $A \vdash A$ – while (1.20) can come out false for there are impossible worlds where Hespeus is not Phosphorus while, all the same, Hesperus remains self-identical. Let us turn to subjunctive conditionals and in particular counter-possibles, the example here being (1.21): 'If Amy had squared the circle, then Amy would be famous', and (1.22): 'If Bob had squared the circle, then Amy would be famous'. Intutively the former is true while the latter false, and hence the two express distinct propositions. We have seen that neither the SPW theorist nor the structuralist can account for the truth of the former and the falsity of the latter. However, if we allow impossible worlds and use Stalnaker's and Lewis's strategy for dealing with counterfactuals, then we make considerable way. Assume that a counter-possible conditional is true if in all – or in all closest – impossible worlds that the antecedent is true, the consequent is also true. Then, assuming that, given an impossible world w in which Amy squares the circle, out of the impossible worlds in which Amy squares the circle, the closest to w are the worlds on which such a profound mathematical achievement is met by instant star status, (1.21) comes out true. However, (1.22) does not, for there seems no natural connection between Bob squaring the circle and Amy being famous to put the former worlds close to worlds in which the latter is also true.

All in all, extended possible world semantics deliver almost all the finess of grain we set out to seek. I have additionally argued that Jago's account is an improvement on Priest's for at least two reasons: Priest is subject to the problem of bounded rationality, – he, so to speak, fails to place a good lower bound – and secondly, he seems to be offering a technical fix rather than a philosophically plausible and illuminating account of meaning and intentionality when it comes to issues arising from SII. Jago's solution to the problem of bounded rationality is an ingenious one for it does not come across as *ad hoc* as selected worlds strategies do: Jago 'selects' worlds by imposing a structure on the state space and not by imposing closure conditions on worlds. Nevertheless, which worlds lie on the epistemically inaccessible side of the spectrum is a result of what proof system Jago uses to structure his space, thus possibly raising suspicion as to whether there is a covertly *ad hoc* element to his account. Nevertheless, this is no damning criticism, and it certainly remains the preferred extended possible worlds strategy.

There is then the glitch that the most satisfying account – that of Jago's – seems to be committed to conflating certain necessary truths and necessary falsehoods, violating, among other things, intuitions of aboutness. However, as noted earlier, it is arguably the case that such necessary truths and falsehoods are indeed *epistemically on par*, that is, if we are modelling *epistemic* content – as Jago clearly is – then two flat contradictions are epistemically indiscernible in that they both are what no rational agent can have epistemic statuses about, or two obvious truths are epistemically indiscernible in that they both are what all rational agents must believe, know etc. I am not wanting to raise a criticism against Jago in saying that conflating such claims is a 'glitch'; for all Jago cares – epistemic content – this can be a desirable prediction. Nevertheless, the aim of this thesis is to model content *simpliciter*, epistemic and otherwise. And when it comes to judging whether that it is raining and not raining and that this ball is blue and red all over simultaneously, we sure do discern between the two: one is about the weather, the other about a given ball. Aboutness is something the structuralists gave us, but a structured approach to content was found wanting. In the chapters to follow I look into non-structured accounts of meaning that nevertheless attempt to tell us what the proposition a given sentence expresses is about.

It should become more and more clear throughout the thesis, but a certain pattern is already emerging. We seem to have identified two things that we require a satisfactorily hyperintensional account of meaning to account for. Firstly, accommodating the impossible seems vital in delivering 'strong' hyperintensionality – as in, in distinguishing between distinct impossibilities, and thus in accounting for counter-possible conditionals too. Secondly, accommodating for aboutness seems just as important in distinguishing between certain intensionally equivalent propositions, which is something that becomes clear within an intensional setting as soon as we lose the aboutness-respecting results that the absolutely unrestricted totality of worlds delivers by default – by distinguishing any two numerically distinct propositions whatsoever.

3.4.2 The Metaphysics of Impossible Worlds

As promised, and for the sake of completeness, a short discussion of selected topics in the metaphysics of impossible worlds follows. As should have become clear from the introduction already, the metaphysics of impossible worlds do not much concern the current project so the discussion here is kept short. However, I want to at least acknowledge the fact that some may be reluctant to endorse extended possible world approaches precisely because they take issue with the metaphysics that substantiates them. I would find this move largely unconvincing; in our post-Quinean intensional era the heuristic value of possible worlds has convinced most to play along albeit their metaphysics remaining debatable, and I see not why impossible worlds should be treated any differently.

Graham Priest in Priest 1997 defends the following thesis, later dubbed the 'Parity Thesis': impossible worlds should inherit the ontological status of possible worlds. However, we cannot dispense of the metaphysical question quite that easily. This is because the parity thesis has been challenged; genuine modal realism does not seem to be a coherent metaphysical stance towards impossible worlds. As far as I know, only Yagisawa (Yagisawa 1988) has argued that there are concrete, spatio-temporally isolated worlds ontologically on par with the actual world that instantiate impossibilities. Nevertheless, it is important to note why such a stance is, other than very strong – claiming that there are contradictions 'out there' in reality – problematic in an additional sense. Suppose we ascribe to the Lewisian picture of impossible worlds. Then, our saying that it is impossible that $A \wedge \neg A$ 'translates' into the following: there is an impossible world w such that $w \Vdash A \wedge \neg A$. Therefore, the argument goes, given that for Lewis 'at w' is an extensional restrictor of quantification over worlds, it passes through the truth-functional connectives yielding that $w \Vdash A$ and that it is not the case that $w \Vdash A$, which is a contradiction at the actual world.

Ersatz accounts of impossible worlds, representing rather than instantiating impossibilities, do not face this issue, neither do they postulate that there are inconsistencies 'out there' in reality. Moreover, the extension of most ersatz account to impossible worlds is easy and comes at no ontological cost. Consider linguistic ersatzism according to which possible worlds are sets of sentences from a world-making language, and thus according to which worlds represent states of affairs just as language does. To deliver impossible worlds we need only relax the constrains on what sets of sentences count as worlds: allowing for inconsistent or incomplete sets or sentences will take us a long way. However, just as is the case with ersatzism about possible worlds, such theories have faced a number of objections due to Lewis, most prominent amongst them that any ersatz account of worlds needs to eventually resort to some primitive intensional entity, or some primitive intensional notion. However, even just for possible worlds, many have found it preferable to allow for primitive modality rather than face the 'incredulous stare' objection to genuine modal realism.

Jago is a linguistic ersatzist with respect to both possible and impossible worlds. That is, he holds that possible and impossible worlds are metaphysically on par, and that all worlds are constructed from actual entities, but somehow manage to represent non-actual entities and states of affairs. The actual entities worlds are constructed out of are for Jago sentences of a world-making language. But central to any ersatz account is addressing the question of how such worlds represent particulars, properties and states of affairs, and to answer that we must look closer at our world-making language and how its names represent particulars, its predicates properties, and its sentences states of affairs. Linguistic ersatzism has it the easiest out of all the ersatz account of worlds, for the obvious answer is that worlds being sets of sentences represent just as ordinary sentences do. But this is not the case of an ordinary language such as English or Greek; Jago invokes a special, world-making language and we must look closer at its constituents and how they represent.

But to do that we must first take a detour since Jago holds an uncommon view of properties: he argues that (1) there are negative properties, and at least some of them are independent of any positive property, and (2) that properties are abstractions from facts. Just as positive properties are ways things are, negative properties are ways things are not. Just as I am brunette and hence possess the property of brunette-ness, I am not blond and therefore possess the property of non-blondness. The property of non-blondness might be a poor example for we may set-theoretically construct it taking the complement of the property of blondness, but it is central to the theory that there are negative properties that are ontologically independent of any positive property whatsoever.[11]

Secondly, Jago does not view properties as primitive universals or as tropes but instead as abstractions from worldly facts. To be perfectly clear, that means that Jago ascribes to an ontology of worldly facts – both positive and negateve – where a fact is an entity that is neither a particular, nor a property but a particular-possessing-a-property. Moreover, for Jago facts are substantial non-linguistic entities – hence 'worldly' – and they are primitive and unstructured – or at least those that are logically atomic are – and the totality of them is the world. Now, given an ontology of such facts, we may answer the question of whether an object o possesses a property P – it does so long as $[Po]$ exists – and the question of what makes sentence Po true – the fact that $[Po]$ – and hence facts are truthmakers for sentences. Properties then are *abstractions* from facts, where the notion of abstraction is a primitive for the theory. Albeit a primitive, property-abstraction can be modelled as an operation as follows: given a fact, say $[Po]$, we abstract o away, which leaves as with the property P. Properties then can be modelled as functions from particulars to facts. On this understanding, properties are, albeit entities on their own right, dependent or else unsaturated entities; for a property P to exist some particular must exemplify it, since for a property F to exist it must be an abstraction from some fact $[Po]$. Therefore, for Jago, there are no uninstantiated properties.

It remains to see how particulars are represented by the world-making language. Jago proposes that objects – actually existing, merely possible ones, and impossible ones too – are represented using property-bundles. However, he does not wish such property bundles to behave like descriptions i.e. non-rigidly and coffering descriptive content to their denotations. Instead he wants property-bundles to be directly referential and rigid, to denote as names and descriptions do. He does so as follows: take the property bundle $b = \{P_1, ..., P_n\}$, it behaves like Dthat(the $P_1 \wedge ... \wedge P_n$). 'Dthat' is Kaplan's famous device from (Kaplan 1978), which attaches to a definite description to form a complex that picks out an object rigidly – the unique object that is $P_1 \wedge .. \wedge P_n$ – but does not attribute the property of being P_i for any $i \in [1, n]$ to the object. Hence, property bundles like b do behave like names *à la* Kripke, and Jago calls them 'bundle-names'. But the issue of representation of mere possibilia is not yet close, for given Jago's actualist assumptions, there are no mere possibilia, and hence bundle-names are empty. Jago then argues that they nevertheless have representational content although we may rightfully only attribute to them when they combine with predicates of our world-making language. So, for example, assume that $b_{shoes} = \{$.., are red, are shoes, belong to me, ..$\}$ is a name-bundle that combines with predicates to write world-making sentences that represent situations in which I have a particular pair of red shoes. In this sense, b_{shoes} has representational powers, it is vital to construct sentences that represent that some red shoes of mine are such-and-such, which is what should be understood when Jago says that a bundle-name represents a mere possible individual.

[11] This is an important point for much of Jago's motivation comes from wanting to overcome the problem of alien properties, and if all negative properties where simply negated positive properties he could not deliver the alien properties that he wants.

This is then how sentences of the world-making language represent. Jago's ersatz worlds are simply sets of world-making sentences that represent by inclusion: for a world-making sentence 'A', w represents that A if and only if '$A' \in w$. At this stage, and to allow for flexibility, any non-empty set of world-making sentences counts as a world, be it incomplete or inconsistent. Given a totality of worlds thus delivered, it is not straightforward when a world counts as possible or impossible. Jago does not attempt to tackle this issue, noting – among other possible answers – that one can simply be a deflationist about necessity and insist that there is nothing metaphysically special about possible worlds other than that in a given context we consider that the world represents a metaphysical possibility. The current is a particularly short summary of the discussion to be found in chapters 4 and 5 of Jago 2014, and as such many an interesting points have been left out, black-boxed or underdeveloped. I could not have hoped for a comprehensive discussion of all the issues Jago raises, but the current should suffice to metaphysically substantiate the framework Jago builds to account for content – epistemic included – on the face of the problem of bounded rationality.

This section is admittedly dismissive of many a topics in metaphysics. Unfortunately, I cannot afford to consider all interesting and somewhat relevant topics that I touch upon. I included this section for three reasons: firstly, I did not wish to be completely quietist, but more importantly, I wanted to remind the reader that the main focus of the thesis is on hyperintensionality, and it is considerations of finess of grain that I take to be most decisive in navigating through the literature I have chosen to consider.

Chapter 4

Yablo

4.1 Introduction – Outline

In his 2014 book *Aboutness* (Yablo 2014a) – as supplemented by Yablo 2014b – Stephen Yablo develops an original and sophisticated account of subject matter based on standard possible worlds semantics, and discusses a plethora of applications, ranging from capturing partial truth as truth of a part, to providing an incremental account of some indicative conditionals.[1] Intriguing as all those endeavours are, they are not the focus of the current chapter. Currently, I am interested in presenting the basic framework Yablo operates within, starting from his notions of truthmakers and falsemakers for sentences and propositions, using those to construct subject matters, and finally arriving at his notion of a 'thick' proposition, so to ask: how finely do Yablo's propositions cut? Yablo's account is extremely relevant to my project for at least the following reasons: it is an intensional account of meaning that, as will soon become evident, can draw many a hyperintensional distinctions, it is a non-structured account of meaning that attempts to capture intuitions of aboutness, and it introduces us to the notion of a truth/falsemaker which makes for a smooth passing to the work of the last theorist I examine, Kit Fine. The chapter closes with a suggestion: combining Yablo's work with Jago's extended possible world semantics, the reasons for which become evident when examining Yablo's account against the data.

4.2 Thick Propositions

4.2.1 What they are

No long-winded introduction is necessary, for as stressed above, Yablo's account is based on the familiar SPW framework. According to Yablo, a sentence's X *intensional content* is its SPW content: the set of possible worlds in which it is true, or the characteristic function thereof. Written $|X|$, Yablo calls this the *thin* proposition X expresses. But for every sentence X there is associated with it a *thick* proposition: this is the *directed content* of X, written $\mathbf{X}$ and comprising of the intensional content of X, together with X's subject matter. So all that remains to be seen is what the subject matter of X is. This is no small thing and I start by examining what object a Yabloesque subject matter is, and then turn to sentential subject matters.

[1] For the latter, see also Yablo 2016.

According to Yablo, a subject matter is a *similarity* relation on the set of possible worlds W, where a similarity relation R on a set S, is a relation $R \subseteq S \times S$ that is reflexive and symmetric.[2] A similarity relation R on S induces a *division* of S, much in the same way that an equivalence relation on a set induces a partition of the set. A division of S is a collection D of subsets of S such that the following conditions are satisfied:

1. The members of D – referred to as 'cells' – are incomparable:
$$(\forall c, c' \in D)(c \not\subset c' \wedge c' \not\subset c).$$
2. The cells are jointly exhaustive of S:
$$\bigcup D = S.$$
3. Each cell is closed under the similarity relation:
$$(\forall c \in D)(x \in c \rightarrow (\forall y \in S)((x, y) \in R \rightarrow y \in c)).$$
4. Each set closed under R is a member of D:
$$(\forall s \subseteq S)((x \in s \rightarrow (\forall y \in S)((x, y) \in R \rightarrow y \in s)) \rightarrow s \in D).$$

Or more succinctly, we have that the division D of S corresponding to a similarity relation R is
$$\{c \mid c \text{ is maximal among subsets } s \text{ of } S \text{ such that } (\forall x, y \in s)((x, y) \in R)\}.$$
Similarity relations determine divisions and *vice versa*, in that we have the following equivalence:
$$(x, y) \in R \iff (\exists c \in D)(x \in c \wedge y \in c).$$
Mathematically speaking this construction is straightforward.[3] But what is the motivation behind taking divisions of W or equivalently similarity relations on W to model subject matters?

Let us take an example. Consider **the number of planets**, construed as a subject matter in that, intuitively, "There are 8 planets" says something about **the number of planets** while "My dog is barking" does not. We may view **the number of planets** as a similarity relation R on W, in that for any $w, w' \in W$, $(w, w') \in R$ if and only if the planet-count of w is equal to the planet-count of w'. Equivalently, we may see it as the division D of W induced by R, each of which cells contain worlds that are equinumerous planets-wise. The reader may have noticed that **the number of planets** construed this way is not only a similarity relation, but an equivalence relation on W since it satisfies transitivity: if w has the same number of planets as w' does, and if w' has the same number of planets as w'' does, then w and w'' have the same numbers of stars. However, Yablo holds that to insist on transitivity as a general condition subject matters should satisfy is too restricting, for consider the following subject matter: **the number of planets give or take one**. If w has 8 planets, and w' has 9 planets, then w, w' are **the number of planets give or take one**-similar. And if w'' has 10 planets then w', w'' are **the number of planets give or take**

2

Reflexivity: $(\forall x \in S)((x, x) \in R)$.

Symmetry: $(\forall x, y \in S)((x, y) \in R \rightarrow (y, x) \in R)$.

3 Refer to Yablo 2014b for more details, and to Hazen and Humberstone 2004. In the latter, what we here call 'divisions' – following Yablo – are called 'decompositions'.

one-similar. But it is not the case that w, w'' are **the number of planets give or take one**-similar, *contra* transitivity.[4]

What is more, there is an intuitive sense in which certain subject matters include others, and such intuitions can be captured within the current set up. Consider for example **the number of planets** and **the number of planets and their sizes**. Intuitively, the latter includes the former, for to say something about the number of planets and their sizes, is to say something about the number of planets. It is easy to make this precise: consider two divisions D, D' of W. We say that D *includes/ refines* D' if and only if each cell c' of D' contains a cell c of D. So, for example, we have that **the number of planets and their sizes** includes **the number of planets**, because each way for things to be with respect to **the number of planets and their sizes** – a *way* being captured by a cell of the division – implies a way for things to be with respect to **the number of planets**. So we now know what a subject matter is as an object in and of itself, and how subject matters may relate to one another. However, the question we set out to answer was, given a sentence, what is its subject matter?

An initial suggestion, a version of which is championed by Lewis 1988, is that a sentence X is wholly about a subject matter $\mathbf{x}$ just in case X's truth value supervenes on which cell of $\mathbf{x}$ we are in, that is, if and only if for any $w, w' \in W$, if $(w, w') \in \mathbf{x}$, then X has the same truth-value in w and w'. This is a good start but it does not answer our question, for there are many subject matters a given a sentence X is about in this sense. For example, "The number of planets is 8" is wholly about **whether the number of planets is 8 or not** and **the number of planets**. Taking a cue from the work of the inquisitive semantics group of Amsterdam, see Ciardelli, Groenendijk, and Roelofsen 2013 for an introduction, Yablo's divisions may be seen as questions, the former corresponding to "Is the number of planets 8?" – resulting in two cells, one comprising of worlds with 8 planets, and one of worlds with a number of planets other than 8 – and the latter to "What is the number of planets?" – resulting in a partition of W into cells each of which contain worlds with an equal number of planets, that is a 0-planets worlds cell, a 1-planets worlds cell, a 2-planets worlds cell, etc. Yablo's suggestion is then that *the* subject matter of a sentence X is the division of $|X|$ corresponding to the question "Why is X true?", and the anti subject matter of X the division of $|\neg X|$ corresponding to "Why is X false?", and the overall subject matter of $|X|$ is its subject matter and subject anti-matter taken together.

Let us focus on subject matter and examine how Yablo motivates his suggestion. Subject matter inclusion gives us a way to order the subject matters that X is wholly about, in Lewis's sense. The weakest such subject matter is always **whether** X **or not**, but it cannot be the subject matter of X for subject matter is not only concerned with truth, but with reasons for being true. The strongest subject matter X is wholly about is the trivial partition of the space corresponding to **how things stand with respect to everything**. Again, this cannot be the subject matter of X for subject matter is concerned with reasons for X being true and nothing else. These considerations give us an upper and lower bound respectively, and as soon as we equate those we get that the subject matter of X which is concerned with reasons for the truth value of X – nothing more, nothing else – is the division of the space corresponding to the question "Why is X true?".

Given a sentence X and its subject matter $\mathbf{x}$ so acquired, the cells of the division corresponding to $\mathbf{x}$ are called *the truthmakers of* X. That is, the truthmakers of X are sets of worlds, i.e. (thin) propositions, corresponding to different answers to the question "Why is X true?". But we may now ask, what properties should we expect truthmakers to have given this setting? Let us take

[4]Lewis in Lewis 1988 develops an account of subject matter as equivalence, but Yablo argues extensively in favour of similarity in §2 of Yablo 2014a. In a footnote in p.37 of Yablo 2014a, he admits he would ultimately want to depart even further from Lewis, and consider *covers* of W rather than divisions. Covers effectively drop condition 1. from the definition of division above, and would be better suited for subject matters such as **the approximate number of planets**.

a world w. A fact obtaining at w is a proposition which is true at w, that is, a set of worlds w belongs to. The subject matter $\mathbf{x}$ of X can be seen as assigning to each world w a set of facts that are the truthmakers of X in w. A minimal requirement then for a truthmaker of X in w is that it be a fact that obtains in w, i.e. that it be a proposition true at w. What is more, we want a truthmaker for X in w to be a fact necessitating X, that is, we want it to be a proposition that implies that X. This much is clear, but not all facts obtaining at w that necessitate that X are good candidates for truthmakers, for we want it to be the case that a truthmaker for X at w somehow *explains* that X, or equivalently that X holds at *w in virtue of* the truthmaker obtaining at w.

And when we turn to explanation, things get a bit murkier; Yablo certainly has some things to say, but to be sure he offers no complete account of explanation and thus his whole notion of truthmaking is rendered somewhat dubious. Let us suppose that T is a fact obtaining at w and implying that X. According to Yablo, for T to be a truthmaker of X at w "[...] is *something* in the neighbourhood of T effecting an optimal tradeoff between naturalness and proportionality"(Yablo 2014b, p.5), where naturalness and proportionality are analysed as follows. Assume that T_1, T_2 are two facts obtaining at w and implying that X.

Naturalness: T_1 is more natural than T_2 if T_1 is a more *compact* and *principled* set of worlds. For example, that it is raining, is preferred over that it is raining or hailing as a truthmaker for 'It is raining or hailing'. T_1 is also more natural than T_2 if the truthmakers for T_1 are more natural than the truthmakers of T_2 in the above sense. For example, that the chair is empty is preferred over that France has no King as a truthmaker for 'The King of France is not sitting on the chair'.

Proportionality: T_1 is more proportional to X than T_2 if it involves fewer irrelevant extras in whose absence it would still imply X. For example, that it is raining is preferred over that it is 18 degrees Celsius, the wind is blowing at 24 mph and it is raining as a truthmaker for 'It is raining'.

So, all in all, a truthmaker for X at world w is a fact F satisfying the following conditions:

1. F obtains at w, that is $w \in F$;

2. F implies X, that is $F \subseteq X$;

3. F is as natural and proportional to X as any fact satisfying 1 and 2.

Conditions 1 and 2 are intuitive and it is a determinate matter whether they hold or not. As for proportionality, we may take Yablo to be saying that if T_1 and T_2 are equally good candidates for truthmakers of X at w, and T_1 includes T_2, then T_2 is to be preferred – effectively construing his talk of 'irrelevant extras' as hinting at mereological relations induced by subject matter inclusion. Still, I am not certain proportionality is completely uncontroversial, for nothing has been said about the extent to which any two candidates for truthmakers of X at w will be comparable. Even more problematically so, not much has been achieved by way of analysing naturalness. On the intuitive side of things, all that has been said is that certain truthmakers strike us as more natural than others, and the analysis has only brought us as far as compactness and principled-ness, properties of sets of worlds that have themselves been left unanalysed. What is more, it is far from clear what something like an optimal balance between naturalness and proportionality amounts to. Rather unsatisfying as the analysis of the notion of a truthmaker is, I cannot afford to address the matter further. [5]

[5] For a more extensive criticism and an alternative account of truthmakers that attempts to offer an actual analysis of the notion of a truthmaker see Pastan 2016.

Perhaps the analysis is just incomplete, or perhaps Yablo is an instrumentalist about truthmakers, adhering to the idea that we should not bother so much analysing the notion but rather focus on what truthmakers can do for us. So let us recap and turn to examine what truthmakers can do for Yablo on the hyperintensionality front. Given a world w and a sentence X, let us assume we now know what the truthmakers of X at w are. Now gather together the Fs such that F is a truthmaker for X at some w and we have the set of potential truthmakers for X, which just is the subject matter of X as a division of W. Similarly, we get the set of potential falsemakers for X and from these we get the anti-subject matter of X, subject matter and anti subject matter make for overall subject matter, and we finally have the thick proposition that X, which is the thin proposition that X together with the overall subject matter of X.

4.2.2 Recursive Truthmaking

Clearly, thick propositions are more than thin propositions; they are thin propositions coupled with overall subject matter. Hence we can expect that if we take propositions to be thick, then they are individuated more finely than if we take them to be thin. This is the simple, underlying thought behind examining Yablo's account with an eye on hyperintensionality. In this section I zoom in, looking at the data from §1 to see whether thick propositions succeed in making the hyperintensional distinctions simple SPW propositions – thin proposition in Yablo's terms – fail to deliver. However, I cannot really do so until I have a semantics for truthmakers. Yablo does not provide one. He provides two 'pictures', and admits that "[t]he models represent *tendencies* in truthmaker assignment that pull at times in different directions" (Yablo 2014a, p.62). The pictures in question are the *recursive* and the *reductive* approach to truthmakers. Here I only have space for one, and I adopt the recursive approach; Yablo's worries about the recursive approach – discussed in §4 of Yablo 2014a – do not interfere with the data I wish to examine, recursive truthmaking is the more common way to go (see Fine, Jago etc – citation needed), and Fine (find which paper) has argued it is superior to the reductive one.

Before going into the semantics some exegetical points are due. What concerns Yablo is to develop a semantic conception of truthmaking, as contrasted to a metaphysical one. Under the metaphysical conception, truthmaking is seen as a relation between entities in the world – be those metaphysical chunks such as facts and worlds or even objects – and truths. Here we care not for developing a correspondence theory *á la* Russell 1918 or Armstrong 2004, rather we seek an understanding of truthmaking "whereby it can play a foundational role in semantics" (Yablo 2014a, p.54). 'Truthmakers' might thus be a misleading term for Yablo to use, for truthmakers have often been supplied to remedy Tarski-style deflationary accounts of truth, in order to explain how the true sentences of a language are made true by reality.[6] But this is precisely what Yablo does not want to do: Yablo's truthmakers care about what a statement says about reality, but not about whether reality accords with what the statement says. This is unusual, for truthmakers have customarily been used to help with the latter – as for example in correspondence theories –but it is no damning conviction. Perhaps one finds Yablo's theory of truth incomplete, but as MacBride 2016 notes "[..] pointing out that a Tarski-style theory of truth doesn't tell us this, doesn't establish that the theory is lacking in this respect – not unless it has already been established that a theory of truth for a language that fails to explain how its sentences are made true fails to articulate in some critical respect how reality conspires with meaning to deliver their mutual upshot, *viz.* truth."

So far so good, there is nothing the matter with a semantic conception of truthmakers. But Yablo also seems to suggest that we adopt an instrumentalist stance towards facts and truth-

[6]See §3.5 of MacBride 2016.

makers, nicely summarised in the motto "[t]ruthmakers are as truthmakers do [...]" (Yablo 2014a, p.54), whilst requiring that the truthmaking relation be not vertical – between metaphysical entities and truths – but rather horizontal, holding between truths and other truths. And there now seems to be tension between our desiderata: on the one hand, we adopt a purely instrumentalist stance towards truthmakers, on the other we require that they reside exclusively in the semantical realm and that they play a foundational role there. That makes for rather unfounded semantics, and we cannot avoid concluding that the theory is lacking, in that, at the very least, it is underdeveloped.

Without further ado I present the recursive truthmaker approach I will be coupling with Yablo's framework. Most of the clauses for it take their cue from Van Frassen 1969, although the quantifier clause is modified following a suggestion by Yablo.[7] Lastly, it need be noted that Yablo does not endorse the clauses for the atomic sentences, since it is those clauses precisely he finds most objectionable in the recursive picture.[8] Yablo's objections do not come into play in examining the hyperintensional data I am interested in, and I need a full semantics for a first order language in order to examine the data and be consistent with work in previous chapters.

We begin with a first order language $\mathcal{L}$ with identity. The logical symbols of $\mathcal{L}$ comprise of the connectives $\neg, \wedge, \vee$, the quantifier $\forall$ and a logical relation symbol $=$. The non-logical vocabulary includes a stock of variables, a stock of constants, and a stock of n-ary predicates for each n. For s an assignment of values to variables, D a domain of individuals and δ an interpretation, we have that:

- $\delta_s(c) \in D$, for c a constant.
- $\delta_s(c) \in D$, for x a variable.
- $\delta_s(P, w) \subseteq D^n$, for P an n-ary predicate.

For each n-ary predicate P, we add a symbol $\bar{P}$ to our language to symbolise the n-ary predicate X such that $(\forall w \in W)(\delta_s(X, w) = D \setminus \delta_s(P, w))$. The interpretation extends to complex formulas as per usual, and truth is defined as in the SPW framework in the familiar way based on intensions.

However, it is not simply truth/ falsity we are interested in but reasons for being true/ false. Reasons for being true/ false are truthmakers/ false makers, so we now want to find, given a formula ϕ what its truthmakers and falsemakers are, so we may arrive at thick propositions.

Definition 8 (Recursive Truthmakers)**.** On the recursive approach, given a formula ϕ its truthmakers and falsemakers are given by the following simultaneous recursion, where by $TM(\phi)$ I symbolise the set of truthmakers and by $FM(\phi)$ the set of false makers of ϕ:

Atomic–1 $\phi \equiv Pc_1..c_n$

- $TM(\phi) = \{\{|Pc_1...c_n|\}\}$, where $|Pc_1...c_n|$ is the intensional content of ϕ, and thus $\{|Pc_1...c_n|\}$ is the *fact that* ϕ.
- $FM(\phi) = TM(\bar{\phi}) = TM(\bar{P}c_1...c_n)$.[9]

[7] See Yablo 2014a §4.4.

[8] Yablo disagrees with the present approach insofar as it implies that simple sentences are true for simple reasons. Yablo certainly has a point but the matter is irrelevant for the discussion in the following subsection and I thus suppress it. See §4.2 and 4.3 of Yablo 2014a for his objections in detail.

[9] Note that the truthmaker for an atomic sentence, say, Pa, is the atomic fact $\{|Pa|\}$. This contrasts slightly with the informal discussion in the previous section where Yablo argued that a truthmaker is simply a proposition. Given the latter we would expect that the truthmaker of Pa is the thin proposition $|Pa|$, and it for our technical convenience that

Atomic–2 $\phi \equiv c = c'$

- $TM(\phi) = \{\{|c = c'|\}\}$.
- $FM(\phi) = TM(\bar{\phi}) = TM(c \neq c')$.

Complex–1 $\phi \equiv \neg\psi$

- $TM(\phi) = FM(\psi)$.
- $FM(\phi) = TM(\psi)$.

Complex–2 $\phi \equiv \psi \wedge \chi$

- $TM(\phi) = \{\tau \cup \sigma \mid \tau \in TM(\psi), \sigma \in TM(\chi)\}$.
- $FM(\phi) = FM(\psi) \cup FM(\chi)$.

Complex–3 $\phi \equiv \psi \vee \chi$

- $TM(\phi) = TM(\psi) \cup TM(\chi)$.
- $FM(\phi) = \{\tau \cup \sigma \mid \tau \in FM(\psi), \sigma \in FM(\chi)\}$.

But how about quantified truths? Let us concentrate here on universal truths, and treat existential ones simply as their dual. In his *Logical Atomism* lectures, Russell concedes with certainty that there are general facts over and above the atomic ones which make generals truths true. However, he soon after admits: "I do not profess to know what the right analysis of general facts is" (Russell 1972, pp.71–72). Russell's conviction about the existence of general facts is rather irrelevant to the current project; as stressed above, we are after a horizontal, semantic conception of truthmakers and not interested in developing a correspondence theory, grounding truth to a metaphysical reality of facts. For the same reason, the work of D.M.Armstrong Armstrong 2004 is not entirely relevant either. However, Armstrong, contrary to Russell, does attempt an analysis of facts that make general truths true:

> My idea is that the truthmakers for such truths are facts, states of affairs, having the following form: a relation, which I will here call the *Tot* relation, holds between a certain mereological object and a certain property.

The truths Armstrong is alluding to above are universal truths of the form "$a, b, c, ...$ are all the A's", and the idea is that the object – made out of the mereological fusion of $a, b, c, ..$ – *totals* the property A. Regardless of whether one agrees with unrestricted mereological fusion, or whether one agrees that such a fact as $a, b, c, ...$ totalling A is a legitimate, non-negative wordly fact, the hint I take from Armstrong is the following. Assume that $a, b, c, ...$ are the A's. That a is B, and that b is B, and that c is B, ... do not suffice to make the general truth "All A's are B's" true, for in addition we need somehow guarantee that $a, b, c, ..$ are all the A's that there are. For consider the alternative. According to Van Frassen 1969, the clause for the universal quantifier should be as follows:

$$TM(\forall x\phi) = TM(\phi(a) \wedge \phi(b) \wedge \phi(c) \wedge ...)$$

we take the singleton $\{|Pa|\}$ instead. I have aimed at keeping the clause for the truthmakers for conjunction and false makers for disjunction as is usually found in the literature, i.e. a truthmaker for a conjunction is the *union* or *fusion* of some truthmakers of its conjuncts. It would be okay to stick to $|Pa|$ and change the latter clauses but I have chosen the current 'fix'.

However, recall that a truthmaker for $\forall x\phi$ is a fact, a proposition that, among other things, implies $\forall x\phi$. And now consider that we may be working with worlds with variable domains such that in w' there exists an object o that does not exist at w. Then if $\phi(a) \wedge \phi(b) \wedge ...$ is a truthmaker for $\forall x\phi$ at w, then it is not a truthmaker for it at w' for we would additionally need that $\phi(o)$. What we are in need of is a fact that guarantees that $a, b, c, ..$ are all the objects that there are, and of course that they are all ϕ.

An initial suggestion would be the following: $\forall xFx$ is made true by $Fa, Fb, Fc, ...$ combined with the fact that $a, b, c, ...$ are everything. Let us abbreviate the totality fact that $a, b, c, ..$ are everything by T, and for arbitrary A the fact that $a, b, c, ...$ are all the A's by T_A, and recall that according to Yablo there's two sides to truthmaking: necessitation and explanation. And we may ask, is the totality fact explanatorily on par with the set of all instances falling under a universal truth? That is, does T explain that $\forall xFx$, or is it rather that T explains why $Fa, Fb, Fc, ...$ explains and necessitates that $\forall xFx$? It rather seems that to simply union the totality fact with the totality of instances of a universal truth is to "confuse the issue of what the truthmaker is, with the issue of how it acquires that status" (Yablo 2014a, p.63). However, our semantic conception of truthmakers allows us to make sense of the idea of truthmakers for truthmakers, and this seems to be what the totality fact is to the set of $Fa, Fb, Fc, ...$ But not quite, for truthmakers necessitate and it is beyond the current scope to ask whether the totality fact suffices to necessitate that $Fa, Fb, Fc, ..$ make $\forall xFx$ true. Yablo's suggestion is then the following: $\forall xFx$ is made true by $Fa, Fb, Fc, ..$ *qua complete list of instances*. As Yablo notes, this turns the totality fact to something akin to a presupposition: $Fa, Fb, Fc, ...$ *qua everything* implies that a is F for it fails if Fa does. However, it does not imply T, for when T fails $Fa, Fb, Fc, ...$ *qua everything* has a truth value gap. As for the false makers of a universal claim $\forall xFx$, any instance of $\bar{F}x$ will do. So, borrowing notation from the literature on presuppositions, we may state the clauses for the universal quantifier as follows:[10]

Complex–4 $\phi \equiv \forall x\psi$

- $TM(\phi) = TM(\psi(a) \wedge \psi(b) \wedge \psi(c) \wedge ... \wedge \partial T_\psi)$
- $FM(\phi) = FM(\psi(a) \vee \psi(b) \vee \psi(c) \vee ...)$

And dually for the existential quantifier:

Complex–5 $\phi \equiv \exists x\psi$

- $TM(\phi) = FM(\forall x\bar{\psi})$
- $FM(\phi) = TM(\forall x\bar{\psi})$

This is then the truthmaker semantics I will be pairing with Yablo's framework. They are not the semantics Yablo endorses, for Yablo does not endorse a single semantics. Where Yablo would object to the current clauses I have pointed out throughout the exposition, but I believe his worries are largely irrelevant to my current goal, the goal – pursued in the subsequent section – being to examine the data from §1, and find out how finely thick propositions cut. It need be noted that the semantics do not include clauses for the conditional. It would take me too far afield to do any justice to Yablo's work on the subject matter, and those data that pertain to conditionals will be discussed somewhat more informally, yet rather decisively still.

[10]The point in common with presuppositions that suggests ∂ as a good notation here is, very roughly speaking, the following: if X presupposes Y in a certain sense, then, if Y is true then X is subject to certain truth conditions, whilst if Y is false then – *qua* Frege 1892, Strawson 1950 and others –Y we have a presupposition failure and X exhibits a truth value gap. For a longer discussion see §5.2 of Beaver and Bart 2014.

4.3 Conclusion

Allow me to start by noting the rather obvious: thick propositions *à la* Yablo cut at least as finely as standard possible worlds – or else thin – propositions. This can be readily acknowledged: a thick proposition comprises of the corresponding thin proposition, and its subject matter, so if we have two thick propositions say $A = \langle |A|, \sigma(A)\rangle, B = \langle |B|, \sigma(B)\rangle$ – where $|X|$ is the thin proposition that X, and $\sigma(X)$ is the overall subject matter of X – and $|A| \neq |B|$, then A is not the same as B. Hence we concentrate on the data from §1, propositions which are, as has been extensively argued, *wrongly* conflated within the SPW framework.

4.3.1 The Data

To begin with, it has been argued that the identification within the SPW framework of intuitively distinct necessary truths such as $P \vee \neg P$ and $Q \vee \neg Q$ is problematic, and the problem, even on an intuitive level, seems to be that these truths cannot be one and the same for they concern different things – one is about how things stand P-wise, the other how they stand Q-wise. Indeed, given Yablo's framework, we may distinguish between such necessary equivalents. Let us take as an example the propositions expressed by the following two sentences,

Amy is dancing or it is not the case that Amy is dancing.
This ball is red or it is not the case that this ball is red.

which translate into our first order language as

$$Da \vee \neg Da \tag{4.1}$$
$$Rb \vee \neg Rb, \tag{4.2}$$

in the obvious way.

Granted, the thin propositions (4.1) and (4.2) express are one and the same – they correspond to the set of all possible worlds.[11] However, their subject matter differs and to convince one of that it suffices to show that there is a truthmaker for one that is not a truthmaker for the other. From our semantics, the clause for atomic propositions gives us that $\{|Da|\}$ is a truthmaker for Da, and the clause for disjunction that it is therefore a truthmaker for $Da \vee \neg Da$. However, {|Da|} is neither a truthmaker nor a false maker for Br, hence it is not a truthmaker for $Br \vee \neg Br$. The subject matter of (4.1) is therefore different to that of (4.2), so the overall subject matters of the two are distinct, and hence the two express different thick propositions. This is the intuitively right result, and even the reasons why the two are distinguished seem to respect our intuitions: loosely speaking, the propositions expressed by "Amy is dancing or is not dancing" and "This ball is red or is not red" differ because – to take one – the former is concerned with whether Amy is dancing, while the latter is completely indifferent to it.

Similarly we can distinguish between intuitively distinct, necessarily false logically complex propositions such as the ones expressed by the following pair of sentences:

Amy is and is not dancing (4.3)
This ball is and is not red (4.4)

[11] For simplicity, I assume here that atomic propositions are truth evaluable even in worlds where the objects they concern do not exist, as for example would be the case in a constant domain semantics. However this is not a vital assumption: we could have taken the pair of propositions expressed by "Amy is dancing or Amy is not dancing" and "Amy is tall or Amy is not tall" instead, which, even considered as partial propositions are according to the SPW framework, one and the same although one concerns Amy's height while the other her dance floor activities

written as $Da \wedge \neg Da$ and $Rb \wedge \neg Rb$ respectively. Intuitively, the sentences do not say the same thing: one concerns Amy, the other the colour os this ball. However, the thin propositions the sentences express are identical, and equal to $\varnothing$. That is because in no possible world do we have that both Da and $\neg Da$, or that both Rb and $\neg Rb$. But we can discern the propositions expressed by (4.3) and (4.4) based on their subject matter, as above. Our semantic clauses have it that $TM(Da) = \{\{|Da\}\}, TM(\neg Da) = FM(Da) = FM(\bar{D}a) = \{\{|\bar{D}a|\}\}$, and hence that $TM(Da \wedge \neg Da) = \{\tau \cup \sigma \mid \tau \in \{\{|Da|\}\}, \sigma \in \{\{|\bar{D}a|\}\}\} = \{\{|Da|\} \cup \{|\bar{D}a|\}\} = \{|Da|, |\bar{D}a\}$. Similarly we get that $TM(Rb \wedge \neg Rb) = \{|Rb|, |\bar{R}b|\}$, and since the propositions have distinct sets of truthmakers they are distinct propositions – as required.

The above examples show that Yablo's framework is hyperintensional: we have with it a means of distinguishing necessarily equivalent propositions. I have not given a semantics for intentional operators, but it is easy to see that if, say, belief was treated purely relationally as a relation that holds between epistemic agents and thick propositions, then if we modelled it in the standard way as a modal operator, then we would be able to explain why agents may believe some truths but not all their necessary equivalents. But this is as far down the scale as the current framework can take us; the rest of this section is devoted to showing that it fails to deliver all the finess of grain that we are seeking.

To begin with, although some intuitively distinct necessary truths and falsehoods are distinguished, some others – interestingly logically atomic ones – are conflated. For necessary truths here I have in mind variants of Frege's puzzle. Take the propositions expressed by "Hesperus is Hesperus" and "Hesperus is Phosphorus" written as $h = h$ and $h = p$ respectively. It has been argued extensively that these two sentences express distinct propositions, for they may differ in terms of informativeness, and certain epistemic agent may hold varying intentional stances towards them. The thin propositions $|h = h|$ and $|h = p|$ are nevertheless one and the same – W. Can we distinguish between the thick propositions expressed by the two sentences as we previously did, in terms of their subject matters? It seems not. The subject matter of $h = h$ is $\{|h = h|\}$, and its subject anti-matter $\{|h \neq h|\}$. The subject matter of $h = p$ is $\{|h = p|\}$ which is the same as $\{|h = h|\}$, and its anti-matter $\{|h \neq p|\} = \{|h \neq h|\}$. Therefore, the thin proposition expressed by the two and their overall subject matters being identical, the thick propositions expressed by the sentences are identical.

But one might now argue that the fact that Yablo does not endorse the clauses for the atomic propositions stated in definition 8 is far from irrelevant to my current project: it is precisely the identification $TM(c = c) = \{\{|c = c|\}\}$ that I am relying on in the above reasoning. Note however that even if some other proposition, say ϕ were to be a truthmaker for $h = h$, then at the very least it would necessitate $h = h$. And given the set of worlds we are working with is the totality of possible worlds, all $h = h$-worlds are $h = p$-worlds. Therefore, ϕ would necessitate $h = p$ as well. All we can hope for then is that ϕ explains $h = h$ but does not explain $h = p$ – for otherwise ϕ would be a truthmaker for the latter all the same. But can there be a fact construed as a set of possible worlds that explains – in the relevant sense – $h = h$ but not $h = p$? It seems not, but Yablo has not given us nearly enough on explanation to fully address this question. Let us go back to working with an intuitive notion of truthmaking, and assume we have a fact/ proposition ϕ that makes $h = h$ true. Given that Hesperus is numerically identical to Phosphorus, and given that this is necessarily the case, limited to possible worlds as we are, it follows that the fact that Hesperus is Phosphorus is the very same fact as the fact that Hesperus is Hesperus – facts here are construed as sets of possible worlds. Hence if the latter proposition is made true by ϕ, then so is the former by a simple application of Leiniz's Law. The inability of the current framework to differentiate between the two proposition traces back not to our semantic clause affirming that simple truths are true for simple reasons, but to the fact that without impossible worlds that Hesperus is Hesperus is the very same proposition as that

Hesperus is Phosphorus and to require that a fact makes one true but not the other would be to require that a fact both makes and does not make the very same proposition true.

The limitations of the framework due to its inability to deal with the impossible are even more striking when looking at intuitively distinct, logically atomic necessary falsehoods. Let us take the following pair to illustrate the point:

$$\text{Amy squared the circle} \tag{4.5}$$
$$\text{Bob squared the circle} \tag{4.6}$$

which we translate as Sa and Sb, respectively. Again, the thin propositions expressed by (4.5) and (4.6) are one and the same: $|Sa| = |Sb| = \varnothing$ – that is to say, there is no possible world in which Amy squared the circle, and no possible world in which Bob did. However, intuitively, the two sentences do not say the same thing – one is saying something about Amy's mathematical achievements, the other concerns Bob's. We would thus expect to be able to distinguish between the propositions expressed by the sentences based on their subject matters. But we cannot do that. Precisely because there is no *possible* world in which either Amy or Bob square the circle, their subject matters are $\{\{|Sa|\}\} = \{\{\varnothing\}\} = \{\{|Sb|\}\}$, and subject anti-matters $\{\{|\bar{S}a|\}\} = \{\{W\}\} = \{\{|\bar{S}b|\}\}$. Therefore, the thin propositions expressed by the sentences and their overall subject matters being identical, the thick propositions expressed by (4.5) and (4.6) are one and the same. Again, the fact that Amy squared the circle is according to Yablo the very same fact as the fact that Bob squared the circle, and there cannot be any fact *qua* set of possible worlds making one true but not the other. The problem is again not with our semantic clauses for atomic propositions but with the identification of the former fact with the latter, an identification forced upon us as soon as we limit ourselves to possible worlds only.

I have not developed here a semantics for intentional operators or conditionals. However, one's favourite possible world semantics for intentional operators and conditionals can easily work with thick instead of thin propositions. Where purely relational semantics are concerned, we have that when two propositions X and Y are embedded in some context, then the propositions expressed by the sentences embedding X and Y differ if and only if X and Y differ. Therefore, Yablo is able to predict that, say, "Amy knows that $2 + 2 = 4$" may be true while "Amy knows that $e^{i\pi} = -1$" may be false for 5 year old Amy, since he can distinguish between $2 + 2 = 4$ and $e^{i\pi} = -1$ to begin with. However, he must take it upon him that "The ancients knew that Hesperus was Hesperus" has the same truth value as "The ancients knew that Hesperus was Phosphorus", contrary to strong intuitions, precisely because he cannot distinguish between "Hesperus is Hesperus" and "Hesperus is Phosphorus". We encounter the same problem when we turn to conditionals: given that "Amy squared the circle" and "Bob squared the circle" express the same thick proposition, so do "If Amy had squared the circle then Amy would be famous" and "If Bob had squared the circle then Amy would be famous", although the former strikes us as true and the latter as false.

4.3.2 The Verdict

Yablo's is a rather conservative suggestion: he retains the standard possible worlds framework, and builds on it to give life to thick propositions. Thick propositions, being tuples of standard possible world propositions and subject matters modelled as divisions of logical space, are entities at least as fine grained as SPW propositions. Indeed they are more fine grained and succeed in making certain hyperintensional distinctions such as between logically complex necessary equivalents. And this is a great result, especially since Yablo's ontology comprises of good old possible worlds. However, thick propositions do not take us all the way; there are a plethora of data we cannot account for if we adopt Yablo's suggestion. Propositions expressed by necessarily equivalent and logically atomic sentences are conflated, and so are any further embeddings of those, assuming no further action is taken to accommodate for those. I have argued that the root of the problem is not with the semantic clauses for the truthmakers of logically atomic propositions, clauses that Yablo himself refrains from adopting. Even if a logically atomic truth X were true for a complex reason, then a truth Y that is the very same as X could not fail to be made true by the same truthmaker, no matter how simple or complex. This is a serious limitation and it is the identification of X and Y that is the culprit, an identification that can only be blamed on our insistence on possible worlds. There are three options: give up the 'possible' bit and try to see how to combine Yablo's framework with impossible worlds, give up the 'worlds' bit and turn to possible situations or states, or give up both and work with states *simpliciter*. In the following section, I explore the former option, seeing how to combine Yablo's work with extended possible worlds semantics, and in the following chapter I explore the latter option as developed in a series of articles by Kit Fine.

4.4 Yablo and the Impossible

The aim of the current section is to build Yablo-style thick propositions from a larger totality of worlds, one that goes beyond the possible. In §3, I looked closely at both Priest's and Jago's extended possible world semantics. A main difference between these two theorists is that Priest works with a completely unrestricted totality of worlds – a totality of worlds that comprises of all the possible, impossible and open worlds – while Jago, in order to overcome the problem of bounded rationality, rules some of them out – he rules out worlds that realise blatant contradictions, and worlds which are open with respect to some fundamental logical rules such as conjunction elimination. Which of the two strategies is best suited for our current purpose of extending Yablo's work beyond the possible?

The answer comes quite naturally. First of all, I concluded §3 by arguing that Jago's account is more satisfactory than Priest's all things considered: it dispenses with the problem of bounded rationality, and it does not employ a dubious first order semantics. But secondly and perhaps more importantly, thick propositions built on top of a completely unrestricted totality of worlds would not be finer than thin propositions built out of the same totality. Priest can already make all distinctions imaginable: for any two numerically distinct X and Y there are worlds in which X but not Y and *vice versa*, therefore the thin propositions that X and that Y are distinct. This is not to rule out that the notion of subject matter has nothing whatsoever to offer to Priest's work – perhaps it does – but not on the hyperintensionality front.

Lastly, there are positive reasons why attempting to combine Jago's and Yablo's work is exciting and promising: we have seen that Jago conflates distinct contradictions such as $A \wedge \neg A$ and $B \wedge \neg B$, precisely because he rules out all worlds in which either is true. With an eye on hyperintensionality, this is problematic. Jago is explicitly concerned with modelling epistemic content, and, arguably, distinct contradictions are epistemically *on par*. I am interested in modelling content *simpliciter*, and since there is an intuitive difference between say "It is raining and it is not raining" and "This ball is red and is not red", I want a framework that can capture this. As we just saw in the previous section, Yablo, even without impossible worlds of any kind, can do that. On the other hand, Yablo cannot distinguish between necessarily equivalent and logically atomic propositions that are intuitively distinct. But Jago, even without subject matter, can. One hand washes the other.

Defining thick propositions *à la* Yablo on top of Jago's extended possible worlds framework is easy. We start with Jago's epistemic space: a structured totality of worlds comprising of all possible worlds and all impossible ones that are closed under some basic rules of inference and do not realise blatant contradictions. I henceforth call this 'the state space' and write it $\mathcal{S}$ since, contrary to Jago, I am not interested in modelling epistemic content alone. Subject matter, as well as thin and thick propositions are defined in the exact same way as before, when we were working exclusively with possible worlds. Let X be a sentence of our language. The overall subject matter of X, $\sigma(X)$, is the pair of the set of truthmakers and that of falsemakers for X, $\langle TM(X), FM(X)\rangle$, where $TM(X)$ and $FM(X)$ are given by the simultaneous recursion of definition 8, augmented in the following way. The clauses for atomic sentences of definition 8 imply that extensions and anti-extensions of predicates are exclusive and exhaustive of the domain. However, we now want to accommodate for the impossible and this is no good assumption to carry on; we need to treat truth and falsity *on par*. We have already seen how to do this following Priest, and we use the same method here:

- $TM(Pc_1..c_n) = \{\{w \mid w \vDash^+ Pc_1..c_n\}\} = \{\{|Pc_1...c_n|^+\}\}$.
- $FM(Pc_1..c_n) = \{\{w \mid w \vDash^- Pc_1..c_n\}\} = \{\{|Pc_1...c_n|^-\}\}$.
- $TM(c = c') = \{\{|c = c'|^+\}\}$.
- $FM(c = c') = \{\{|c = c'|^-\}\}$.

Given that we need positive and negative content to accommodate this point, we need to take it into consideration when considering the thin proposition X expresses. $|X|$ is now a pair $\langle |X|^+, |X|^-\rangle$, and although this makes for more nitpicky calculations, it makes for no real bloat in our ideology.

There is a peripheral point that requires attention, but this is nothing out of the ordinary when impossible worlds come into the picture. Recall that, for example, Priest, in defining logical validity and logical implication, has to restrict his attention to possible worlds. This is to be expected; impossible worlds are worlds in which, among other things, the laws of logic may fail, and formulae may come apart from their logical implications. We now have in our hands propositions that are concerned not only with truth conditions, but also reasons for truth conditions, and have already discussed a logical relation other than truth-conditional implication that propositions may stand in: subject matters of propositions may stand in relations of inclusion, and in turn, thick propositions may stand in relations of part hood to each other. Central as the metrology of propositions is to Yablo's project, other than making sure that logical implication is restricted to the possible – by requiring that X implies Y if and only if Y is true in every *possible* world in which X is true – we need to make sure that subject matter inclusion is too. Again, this is easy, especially since Yablo's truthmakers are facts *qua* propositions. Let us see how.

Let X and Y be two formulae expressing the thick propositions $\langle |X|, \langle TM(X), FM(X)\rangle\rangle$ and $\langle |Y|, \langle TM(Y), FM(Y)\rangle\rangle$ respectively. Write $\mathcal{P}$ for the subset of $\mathcal{S}$ consisting of all and only the possible worlds. X truth conditionally implies Y if and only $|X|^+ \cap \mathcal{P} \subseteq |Y|^+ \cap \mathcal{P}$.[12] But X does not include Y unless the inference from X to Y is in addition to truth preserving, aboutness preserving too. And the inference from X to Y is aboutness preserving just in case the subject matter of X refines the subject matter of Y, and the subject anti-matter of X refines the subject anti-matter of Y. The subject matter of X refines the subject matter of Y – and here care is needed to cater to the impossible – if and only if each 'restricted' truthmaker τ of X from $\{\tau \cap \mathcal{P} \mid \tau \in TM(X)\}$ implies a 'restricted' truthmaker σ of Y from $\{\sigma \cap \mathcal{P} \mid \sigma \in TM(Y)\}$. Similarly, the subject anti-matter of X refines the subject anti-matter of Y if and only if each 'restricted' falsemaker τ of X from $\{\tau \cap \mathcal{P} \mid \tau \in FM(X)\}$ implies a 'restricted' falsemaker σ of Y from

[12] I assume that in all possible worlds, for any formula $|X|$, $|X|^+ \cap |X|^= \varnothing$ and dealing with positive content alone suffices.

$\{\sigma \cap \mathcal{P} \mid \sigma \in FM(Y)\}$. In case that X truth-conditionally implies Y, X's subject matter refines Y's and X's anti-subject matter refines Y's, we say that X includes Y.

Given our limited focus, implication and parthood will concern us no further, but these notions are of the utmost importance to the framework in and of itself; that they survive the transition to the impossible certainly merits a mention. To discuss the importance of implication would be to preach to the converted, but it should be noted that Yablo is particularly if not centrally interested in exploring partial truth as truth of a part, and it would be disrespectful so to speak to destroy parthood. Back to my current focus, what remains is to examine more closely just how finely thick propositions of this extended possible world semantics cut, and for that we only need the semantic definitions we opened the current section with. As anticipated in the introduction, thick propositions defined on top of Jago's state space cut just fine.

For convenience, let me dub thick propositions *à la* Yablo defined on top of Jago's state space 'extended thick propositions', in lack of better words. Extended thick propositions cut at least as finely as Jago's extended possible world propositions do, and at least as finely as Yablo's thick propositions do. This is because, in bringing together Jago's and Yablo's works, we took away from neither. In particular, let X and Y be two sentences of our language, and let $|X|, |Y| \subseteq \mathcal{S}^2$ be the thin propositions expressed by them. Then if $|X| \neq |Y|$, clearly $\langle |X|, \sigma(X)\rangle \neq \langle |Y|, \sigma(Y)\rangle$ since their first components differ. This suffices to prove the first of the above claims, and for the latter we need only note that if $\langle |X|^+ \cap W, \langle \{\tau \cap W \mid \tau \in \sigma^+(X)\}, \{\tau \cap W \mid \tau \in \sigma^-(X)\}\rangle\rangle \neq \langle |Y|^+ \cap W, \langle \{\tau \cap W \mid \tau \in \sigma^+(Y)\}, \{\tau \cap W \mid \tau \in \sigma^-(Y)\}\rangle\rangle$ then either $|X|^+ \cap W \neq |Y|^+ \cap W$ in which case $|X| \neq |Y|$, or $\{\tau \cap W \mid \tau \in \sigma^+(X)\} \neq \{\tau \cap W \mid \tau \in \sigma^+(Y)\}$ in which case $\sigma^+(X) \neq \sigma^+(Y)$ and hence $\sigma(X) \neq \sigma(Y)$, or finally $\{\tau \cap W \mid \tau \in \sigma^-(X)\} \neq \{\tau \cap W \mid \tau \in \sigma^-(Y)\}$ in which case $\sigma^-(X) \neq \sigma^-(Y)$ and hence $\sigma(X) \neq \sigma(Y)$; in any case, the extended thick propositions $\langle |X|, \sigma(X)\rangle, \langle |Y|, \sigma(Y)\rangle$ expressed by X and Y differ.

In §3.4, I concluded that the only class of data Jago fails to account for satisfactorily is the class of intuitively distinct yet blatant contradictions – where A and B are distinct sentences of our language, the propositions expressed by $A \wedge \neg A$ and $B \wedge \neq B$ are, according to Jago, one and the same. To be sure, this is no failure *per se*, for Jago cares only about the epistemic content of these sentences, and a case can be made that all blatant contradictions are epistemically on par. However, I care about the content of $A \wedge \neg A$ and $B \wedge \neg B$ *simpliciter*, and since there is an intuitive difference between, say, "It is raining and it is not raining" and "This ball is red and is not red" – one is about the weather, the other about sporting equipment! – I am not content with conflating all blatant contradictions, and indeed, the extended thick propositions expressed by the intuitively distinct blatant contradictions are not identical. The proof requires a one line argument at this stage: Yablo does not conflate such propositions, and given that extended thick propositions cut at least as finely as thick propositions, we have our desired conclusion. But let me go through an example, to see just how this works and that the impossible does not interfere with subject matter in a harmful way.

We have

$$\text{Amy is and is not dancing} \tag{4.3}$$

$$\text{This ball is and is not red} \tag{4.4}$$

written as $Da \wedge \neg Da$ and $Rb \wedge \neg Rb$ respectively, just as before. $|Da \wedge \neg Da|$ is identical to $|Rb \wedge \neg Rb|$ since Jago rules out all impossible worlds realising blatant contradictions, however $\sigma(Da \wedge \neg Da)$ and $\sigma(Rb \wedge \neg Rb)$ differ. This is so because, as our semantic clauses have it, $TM(Da) = \{\{|Da|^+\}\}, TM(\neg Da) = \{\{|Da|^-\}\}$, and hence that $TM(Da \wedge \neg Da) = \{\tau \cup \sigma \mid \tau \in \{\{|Da|^+\}\}, \sigma \in \{\{|Da|^-\}\}\} = \{\{|Da|^+\} \cup \{|Da|^-\}\} = \{|Da|^+, |Da|^-\}$. Similarly we get that $TM(Rb \wedge \neg Rb) = \{|Rb|^+, |Rb|^-\}$, and clearly it should be that, at least, $|Da|^+ \neq |Rb|^+$, hence, since the propositions have distinct sets of truthmakers, they are distinct propositions – as required.

Yablo on the other hand had issues with logically atomic yet intuitively distinct necessary falsehoods. Jago had none, and hence we expect the extended thick propositions expressed by such sentences to be distinct. Again, let us go through an example. Consider the pair

$$\text{Amy squared the circle} \tag{4.5}$$

$$\text{Bob squared the circle} \tag{4.6}$$

which we translate as Sa and Sb, respectively. Yablo conflates the two, as we showed above, but for Jago there are admissible worlds which represent that Amy squared the circle while do not represent that Bob did and *vice versa*. That is, we have $w, w' \in \mathcal{S}$ such that $w \in |Sa|^+$ while $w \notin |Sb|^+$, which goes us that $|Sa|^+ \neq |Sb|^+$, which in turn implies that the thin propositions expressed by (4.5) and (4.6) are distinct, and hence the thick propositions expressed by them are too.

Combing Jago's and Yablo's work in the way suggested in the current section gives rise to propositions that possess the finess of grain we set out to seek. On the hyperintensionality front we have a 'check'. However, the account can raise some worries. Firstly, it inherits a point raised against Yablo earlier: Yablo's instrumentalist stance towards truthmakers, coupled with his semantic conception of truthmakers under which they play a foundational role in semantics is problematic. If pressured, I would drop the instrumentalism and try to offer a real analysis of the notions of naturalness and proportionality that Yablo toys with. This is not the place though, and I must conclude that Yablo's account – and thus the Yablo-Jago account – is underdeveloped.

Another worry is that the very criticism I raised against Chalmers' account in 2.3.3 could very well apply here: the account is *ad hoc* and contrived, a conglomeration of disparate theories. I do not think this is the case for the Yablo-Jago account. Combining the two theories subtracts from neither, while one fills in the gaps of the other quite elegantly. The ideology of the account may make it seem a bit contrived – what with working with double contents *and* subject matters – but we should not overlook the fact that the theory has a great deal of the power, applications of subject matter other than towards attaining finess of grain, having hardly had a mention here. And fitting to the pattern, our ontology is a rather conservative one. To add, Yablo's ideology comes with possible worlds, but, in combining it with Jago's work, what we have essentially done is simply change the underlying ontology, move from one set of words to another. The only worry in the vicinity I am sympathetic towards is the following: extended possible world semantics *á la* Priest capture aboutness, insofar as we get no conflations of propositions that violate intuitions of aboutness, simply because for any two numerically distinct propositions whatsoever there is a world that makes one true but not the other. In other words, the mere ontological wealth of extended possible world semantics proper delivers aboutness. Jago is an extended possible world theorist who, for good reason, loses some worlds, and as a result conflates propositions contra intuitions of aboutness. To then mend that by importing the heavy ideology developed by Yablo seems rather *ad hoc*, for Jago has already denied the 'natural' means by which the general framework he operates within accommodates for aboutness.

Kit Fine, to whom we now turn, gets things the other way round and from the very onset promises to dispense with the problems facing Yablo: his stance towards truthmakers, albeit intsrumentalist, does not raise further worries, and moreover, the very nature of the framework he operates within is meant to cater for both aboutness and the impossible.

Chapter 5

Fine

5.1 Introduction – Outline

In a series of articles – starting with the published Fine 2014 and Fine 2016a and others continuously appearing on his Academia page – Kit Fine develops a truthmaker semantics that constitutes not only an alternative to SPW semantics, but also a genuine alternative to Yablo's account. I add stress to the latter comparison because on the purely semantical side – if one were to skip forward to the semantic clauses for the two accounts, or even the semantic notions developed and their behaviours – the differences between the two would perhaps not be immediately noticed. However, the two theorists' answer to the main underlying ontological question sheds light on how inherently distinct the accounts really are. To be more precise, both Yablo and Fine agree that truth-conditions are entities which stand in relations of truthmaking to statements. Adopting this objectual approach leaves some issues wanting, the first and foremost being the question of the nature of truthmakers. It is here that Yablo and Fine most clearly diverge, and most differences between their accounts can be traced back to their respective ontologies.[1] For Yablo, truthmakers are facts construed as sets of possible worlds, while for Fine truth conditions are to be identified, not with whole worlds, but with fact-like entities he calls 'states' that can be seen as making up worlds. Fine jokingly remarks that "[t]he possible worlds approach is fine but for two features: the first is that possible worlds are worlds, i.e. complete rather than partial; and the second is that they are possible (Fine 2015, p.18)."

As we see shortly, Fine works with a domain of both possible and impossible states, and identifies propositions with the tuple of the set of those states verifying, and the set of those states falsifying the statements expressing them. What sets the two theorists apart is precisely that Yablo starts with an intensional framework and tries as best as he can to built a hyperintensional account of propositions out of it, while Fine from the very onset adopts a finer framework akin to that of Barwise and Perry's situation semantics.[2] In the current chapter, I first present Fine's account in some detail and then turn to examine whether it delivers the finess of grain I have been seeking throughout this project. Given that the extended thick propositions account I have extracted from the works of Jago and Yablo combined has been argued the most satisfactory account of content so far – my focus being on hyperintensionality – I inevitably conclude with a comparison between the latter and Fine's account of content.

[1] Fine himself indulges in such a comparison in Fine 2016d

[2] See Barwise and Perry 1983.

5.2 Fine's Truthmaker Semantics

I present a semantics for a propositional language as developed primarily in Fine 2016b, and then take hints from Fine 2016c, so to investigate how the semantics could be extended to cover a first order language. Quantification is not a major concern of mine, and the discussion is kept short – as was for Yablo. Ultimately, I wish to judge Fine's semantics against the hyper-intensional data that have been my guide throughout the thesis, to which end, presenting a full-blown first order semantics is unnecessary; it would take me too far afield given that a semantics is not readily available and many an issues are at large that are nevertheless peripheral to my main focus.

5.2.1 Semantics on a State Space

Fine starts with a domain $\mathcal{P}$ of possible and actual states whose role is akin to that of the domain of possible worlds in SPW semantics. Fine insist that 'state' is, for him, "a term of art and it is used to cover not just states in the ordinary sense of the word but facts, events, conditions or whatever else may legitimately be regarded as a truthmaker (Fine 2015, p.2)." $\mathcal{P}$ then consists of states such as the actual state of this ball being red, the possible state of my being full, given that I am hungry, and the state of the current weather in Amsterdam being both rainy and windy. This conception of states renders them *partial* in a way possible worlds are not: it is not the case that the truth or falsity of any statement will be settled by any state whatsoever, while it is the case it will be settled by any world. For example, the state that it is currently both rainy and windy in Amsterdam does not settle whether this ball is red, while any possible world in which this ball exists settles whether it is red or not.

A natural way for those raised within the possible worlds tradition to think of states is as parts of worlds. By way of an introduction to the notion of a state, characterising states as parts of worlds may be of heuristic value and moreover, when working with extended models, the characterisation can even be formally substantiated. Nevertheless, it is not a characterisation one should read into. The idea Fine is building towards is that if we were to need worlds, then we could have them, for example by defining them as sets of states satisfying certain conditions of completeness, consistency and the like, or as total and consistent states *simpliciter*. However, as will soon become evident, "possible worlds completely drop out of the picture (Fine 2015, p.18)", for we have no need for them.

Be that as it may, already driving the legitimate if misleading conception of worlds as state-conglomerations is an intuition that our set of states $\mathcal{P}$ enjoys a certain mereological structure thet the domain of possible worlds does not, in virtue of the fact that states are partial – the intuition here being that in some sense, the state that it is rainy is Amsterdam is part of the state that it is both rainy and windy in Amsterdam. Indeed, the basis for Fine's semantics is the notion of a *state space* $\mathcal{S}$ which he defines as a tuple $(\mathcal{P}, \sqsubseteq)$ where $\mathcal{P}$ is a non-empty set – to be thought of as our domain of possible states from above – and $\sqsubseteq$ is a partial order on $\mathcal{P}$ – to be thought of as the mereological relation that intuitively holds between certain states – that satisfies the following condition:[3]

Bounded Completeness: Any subset of $\mathcal{P}$ that has an upper bound has a least upper bound.

[3]To remind the reader, a partial order $\subseteq$ on a set X is a binary relation on X that satisfies the following conditions:

Reflexivity: $(\forall x \in X)(x \subseteq x)$;

Anti-symmetry: $(\forall x, y \in X)(x \subseteq y \wedge y \subseteq x \rightarrow x = y)$;

Transitivity: $(\forall x, y, z \in X)(x \subseteq y \wedge y \subseteq z \rightarrow x \subseteq z)$.

Now, an upper bound for a set $T \subseteq \mathcal{P}$ is an $s \in \mathcal{P}$ such that $(\forall t \in T)(t \sqsubseteq s)$, and a least upper bound for T is an upper bound s for T such that for any upper bound s' of T we have that $s \sqsubseteq s'$. Where state s is the least upper bound for a set of states T we say that s is the fusion of the T's and write $s = \sqcup T$. Given that we are working with a set of *possible* states only, it would be unreasonable to assume that the fusion of arbitrary states exist, since for example, any state having as parts the state that this ball is red and the state that this ball is not red would be an inconsistent or else impossible state. However, if there is a possible state s such that it has as parts all states in $T \subseteq \mathcal{P}$ then this is indication enough that the states in T are jointly compatible and thus it is reasonable to assume that such a state as the fusion of T exists in $\mathcal{P}$. This is then why Fine requires that a state space be bounded complete, and does not require that it be complete – that any arbitrary states in $\mathcal{P}$ have a fusion.

Other than the question of the nature of truthmakers, as soon as one adopts an truth conditional, objectual approach towards content there is a related issue to be tackled: what is the relation of truthmaking/ false making that truthmakers/ falsemakers and statements stand in? Fine is interested in *exact* verification and falsification according to which a state "should be *wholly* relevant, and not just relevant in part to the truth or falsity of the statement (Fine 2016b, p.5)." Granted, this is as of yet no real analysis of exact verification, but we return to the issue later, when constructing our notion of a proposition. Leaving the issue aside for now, the state space and flavour of verification/ falsification being in place, we can already state the semantic clauses for a propositional language $\mathcal{L}$ the formulas of which are constructed from a countably infinite set P of sentence letters $p, q, r, ...$ and the connectives $\neg, \wedge, \vee$.

Definition 9 (Exact Verification/ Falsification)**.** Given a valuation function

$$\begin{aligned} [\cdot] : P &\to \mathcal{P}^2 \\ x &\mapsto ([x]^+, [x]^-) \end{aligned}$$

taking propositional letters to their exact verification-falsification conditions – the pair of sets of states consisting of the set of all their possible verifiers and the set of all their possible falsifiers – we may recursively define what it is for a state s to exactly verify/ falsify a sentence A of $\mathcal{L}$, written $s \vDash^+ / \vDash^- A$:

Atomic $A \in P$

- $s \vDash^+ A \iff s \in [A]^+$
- $s \vDash^- A \iff s \in [A]^-$

Negation $A \equiv \neg B$

- $s \vDash^+ A \iff s \vDash^- B$
- $s \vDash^- A \iff s \vDash^+ B$

Conjunction $A \equiv B \wedge C$

- $s \vDash^+ A \iff (\exists t, u \in \mathcal{P})(t \vDash^+ B \wedge u \vDash^+ C \wedge s = t \sqcup u)^*$
- $s \vDash^- A \iff s \vDash^- B \vee s \vDash^- C$

Disjunction $A \equiv B \vee C$

- $s \vDash^+ A \iff s \vDash^+ B \vee s \vDash^+ C$

- $s \vDash^- A \iff (\exists t, u \in \mathcal{P})(t \vDash^- B \wedge u \vDash^- C \wedge s = t \sqcup u)^*$

Note that the clauses marked with a $*$ for conjunction and disjunction place an existential requirement on $\mathcal{P}$ which is not guaranteed to be met. [4]

5.2.2 Semantics on an Extended State Space

The present semantics is already more fine-grained than SPW semantics – for example, if entailment is defined as A entails B if and only if $(\forall s \in [A]^+)(s \in [B]^+)$, then $A \wedge B$ does not entail A – but we have learned our lesson and wish to extend the current to accommodate for the impossible before we check whether it accomplishes the finess of grain we seek. The first step towards a finean truth maker semantics that accommodates for the impossible is extending the state space to include impossible states, and then, independently, seeing how these states can be put to semantic use. The former, ontological task can, according to Fine, be accomplished in a way as natural and unproblematic as filling out the gaps in arithmetical ordering the rationals leave with irrational numbers to get to the reals. The main idea behind the analogy here is that there are certain gaps in our mereological ordering of the state space – left by subsets of $\mathcal{P}$ which have no upper bound – which can be filled by impossible states, keeping our mereological principles undisturbed. Just as is the case with irrational and rational numbers, impossible states are to be identified on the basis of how they relate to possible states. More precisely, Fine's idea is that impossible states correspond to 'fusions' of possible states such as the state that this ball is red and the state that this ball is not red, where their fusion does not exist. But this is not much progress.

In deciding on what 'virtual fusions' really are, or else in deciding on the entities impossible states are to be identified with, we make use of the notion of an ideal:

Definition 10 (Ideal)**.** A set of possible states $I \subseteq \mathcal{P}$ is an ideal if it satisfies the following two conditions:

Upward Closure: $(\forall p, r \in I)(p \sqcup r \in \mathcal{P} \to p \sqcup r \in I)$;

Downward Closure: $(\forall p \in I)(s \sqsubseteq p \to s \in I)$.

So an ideal is a set of possible states that contains all fusions of its member states, and all parts of its member states. Now, some ideals will contain a member of which all other members are parts in which case we call both said member and said ideal 'principal'. Impossible states are simply identified with ideals that do not contain a principal member. But let me expand on why Fine decides on such an identification.

To begin with, Fine's idea is that impossible states correspond to the 'virtual fusion' of the member states of such sets as they have no fusion in $\mathcal{P}$. The most naive way to turn such a correspondence into precise identification would be to identify impossible states with the the subsets of $\mathcal{P}$ that do not have an upper bound themselves. This is on the right track but faces two challenges. Let s be the state that it is rainy and windy, and t the state that it is not windy, s' the state that it is rainy and t' the state that it is windy. Firstly, both $\{s, t\}$ and $\{s', t', t\}$ seem to correspond to the same impossible state, and secondly, both $\{s, t\}$ and $\{s', s, t\}$ seem to correspond to the same impossible state, but if we were to stick to our naive identification above, we would have two distinct impossible states instead of one in either case. Upward closure takes care of the former, downward closure takes care of the latter, and Fine reassures

[4] Note that the semantic clauses here are very much in the same spirit as Van Frassen 1969 and Yablo 2014a. However, such similarities should not take us aback, since the fundamental differences between Yablo's and Fine's truthmaker semantics have already been discussed in the opening paragraphs of the current section.

us that "[i]t turns out that the two problems above are essentially the only problems that can arise in identifying impossible states with sets of possible states (Fine 2016b, p.6)", hence we need look no further than non-principal ideals.

We arrive at our extended domain of states E by postulating into existence the fusions of non-principal ideals and define our extended state space $\mathcal{E}$ as the triplet $(\mathcal{P}, E, \sqsubseteq)$, where $\mathcal{P}$ is our domain of possible states from above, E is our extended domain, and $\sqsubseteq$ the mereological ordering on E abiding to the very same principles as before. On the semantic side of things, the extension is completely straightforward since we keep the clauses of definition 9 unchanged, but for the clauses marked with $*$ which become $s \vDash^+ A \iff (\exists t, u \in E)(t \vDash^+ B \wedge u \vDash^+ C \wedge s = t \sqcup u)$, and $s \vDash^- A \iff (\exists t, u \in E)(t \vDash^- B \wedge u \vDash^- C \wedge s = t \sqcup u)$, where the arbitrary fusions are now guaranteed to exist.

Now that we have the extended state space in place, it is time to turn to our semantic notions and I concentrate on that of a proposition for obvious reasons. There are many an options of how to define propositions on $\mathcal{E}$ and most of them deliver propositions of distinct finess of grain.[5] Hence, since I will be judging Fine on how many hyperintensional distinctions his account can draw, I need to be particularly careful on this point.

The first choice that need be made is whether we want a unilateral or a bilateral account of content. The simplest option would be to identify a proposition with the set of the states that verify it – in a move familiar from SPW semantics – which would deliver unilateral contents. However, as anticipated from the *simultaneous* recursion of definition 9, I here opt for bilateral contents and identify the proposition that X with the pair $([X]^+, [X]^-)$, where $[X]^+$ is the set of verifiers of X – $[X]^+ = \{s \in E \mid s \vDash^+ X\}$ – and $[X]^-$ the set of its falsifiers – $[X]^- = \{s \in E \mid s \vDash^- X\}$. Given our detailed discussion of impossible world semantics, this move should not be too unfamiliar either. As soon as the impossible comes into the picture treating truthmaking and falsemaking *on par* greatly simplifies things, especially when dealing with negation. And, regardless, bilateral propositions are in principle more fine-grained entities since it could be the case that two propositions share a set of verifiers but not a set of falsifiers, meaning the choice of bilateral over unilateral content is in good spirit; to have imposed a unilateral content on Fine and then criticised him for not delivering sufficient finess of grain would have been rather uncharitable.

As soon as we identify propositions with pairs of sets of verifiers and falsifiers we may wish to impose certain conditions on how such positive and negative contents are allowed to relate. Let us say that two states $s, t \in E$ are compatible if their fusion $s \sqcup t$ is a possible state, and incompatible otherwise. If we want that a proposition that X is not both true and false then we require that no verifier is compatible with a falsifier – $(\forall s \in [X]^+)(\forall t \in [X]^-)(s \sqcup t \in E \setminus \mathcal{P})$ – and if we want that a proposition be either true or false, then we require that any possible state is compatible with some verifier or some falsifier of X – $(\forall s \in \mathcal{P})(\exists t \in [X]^+ \cup [X]^-)(s \sqcup t \in \mathcal{P})$. Imposing these requirements would give us bivalent propositions, and dropping them gluts and gaps respectively; here I do not take a stance – to each their own.

But this is not all the choice there is, for Fine explicitly discusses certain conditions that could be imposed on sets of verifiers/ falsifiers so that pairs of such would qualify for propositionhood. The three he focuses on are the following:

Verifiability/ Falsifiability Every proposition has a verifier/ falsifier.

Closure The fusion of some verifiers/ falsifiers of a proposition is also a verifier/ falsifier of the proposition.

[5] For a more detailed discussion than the one that follows, see §Propositions in Fine 2015.

Convexity Any state t such that $s \sqsubseteq t \sqsubseteq r$ for s, r verifiers/ falsifiers of some proposition, is a verifier/ falsifier of the proposition.

As Fine notes, "[w]e obtain different conceptions of exact verification and different domains of propositions depending upon which of these various conditions we accept (Fine 2015, p.4)." The most liberal approach is to impose no requirements, which delivers the greatest finess of grain, and since my focus is on hyperintensionality, for that reason alone, this is the approach I will be adopting for now. In the following section we see that there are hyperintensionality-related issues – namely that the grains may be too fine – to constraint the domain, and allow me to already note that imposing all three constraints results in a very elegant theory of content one can find discussed and developed in Fine 2015.[6]

I now turn to examine how the semantics could be extended to a first order language $\mathcal{L}^+$ with identity. As before, the logical symbols of $\mathcal{L}^+$ comprise of the connectives $\neg, \wedge, \vee$, the quantifiers $\forall, \exists$ and a logical relation symbol $=$. The non-logical vocabulary includes a stock of variables, a stock of constants, and a stock of n-ary predicates for each n. A model for $\mathcal{L}^+$ is now a tuple $(\mathcal{P}, E, \sqsubseteq, D, \delta)$, where $(\mathcal{P}, E, \sqsubseteq)$ is an extended state space, D a non-empty set which is to be thought of as a domain of individuals and δ an interpretation function such that for an assignment s of values from D to variables we have that:

- $\delta_s(c) \in D$, for c a constant.
- $\delta_s(x) \in D$, for x a variable.
- $\delta_s(P, a_1, ..., a_n) \subseteq E^2$, for P an n-ary predicate and $a_1, ..., a_n$ constants or variables, such that, intuitively, the first member of the pair is the set of states which verify P of $\delta_s(a_1), ..., \delta_s(a_n)$ and the second member the set of states which falsify it.
- As a special case, $\delta_s(=, a, b) \subseteq E^2$ for a, b constants or variables, where the first member of the pair is the set of states which verify $\delta_s(a) = \delta_s(b)$ and the second the set of states which falsify it.

We then have the following clauses for atomic formulas:

$$s \vDash^+ Pa_1...a_n \iff s \in [Pa_1...a_n]^+$$
$$s \vDash^- Pa_1...a_n \iff s \in [Pa_1...a_n]^-,$$

and

$$s \vDash^+ a = b \iff s \in [a = b]^+$$
$$s \vDash^- a = b \iff s \in [a = b]^-.$$

For the quantifiers, Fine proceeds in much the same spirit as Yablo, attempting to reduce quantificational statements to their corresponding truth functional formulae. For example, $\forall x \phi(x)$ reduces to $\phi(a_1) \wedge \phi(a_2) \wedge ...$, where the $\delta_s(a_i)$'s pick out all individuals in D, and similarly $\exists x \phi(x)$ reduces to $\phi(a_1) \vee \phi(a_2) \vee ....$ If this is so and the content of a quantificational

[6]Note that here the issue of whether to impose any constraints is not completely independent of how the semantic clauses are stated. The semantic clause of definition 9 are those compatible with the present most liberal approach, so in a sense the choice here has been anticipated. For how to alter the semantic clauses in the case of wanting to impose these constraints see §The Boolean Operations of Fine 2015.

statement is taken to be the same as the content of its corresponding truth-functional formula, the semantic clause for the quantifiers are simply:

$$
\begin{aligned}
s \vDash^+ \forall x\phi(x) &\iff (\exists s_1, s_2, ... \in E)(s_1 \vDash^+ \phi(a_1) \wedge s_2 \vDash^+ \phi(a_2) \wedge ... \wedge s = s_1 \sqcup s_2 \sqcup ...)\\
s \vDash^- \forall x\phi(x) &\iff (\exists a \in D)(s \vDash^- \phi(a)),
\end{aligned}
$$

and

$$
\begin{aligned}
s \vDash^+ \forall x\phi(x) &\iff (\exists a \in D)(s \vDash^+ \phi(a))\\
s \vDash^- \forall x\phi(x) &\iff (\exists s_1, s_2, ... \in E)(s_1 \vDash^- \phi(a_1) \wedge s_2 \vDash^- \phi(a_2) \wedge ... \wedge s = s_1 \sqcup s_2 \sqcup ...).
\end{aligned}
$$

However, this approach faces the same issue that became clear to the logical atomists already, and drove Yablo to consider a totality claim – and consider it as a presupposition for that matter – namely, that if the domain of individuals that there are does not remain constant across states, then if it possible that there are more individuals than there actually are, then $\bigwedge \phi(a_i)$ for all the a_i's that cover the domain of actual individuals does not suffice to make true the universal claim $\forall x\phi(x)$. Fine also proposes a solution using totality states, that runs along the following lines. For each $A \subseteq D$, there exists a totality state τ_A to the effect that the individuals in A are all and only the individuals that there are. Then the clauses for the universal quantifier become:

$$
\begin{aligned}
&s \vDash^+ \forall x\phi(x) \iff\\
&(\exists s_1, s_2, ... \in E \wedge \exists A = \{a_1, a_2, ...\} \subseteq D)(s_1 \vDash^+ \phi(a_1) \wedge s_2 \vDash^+ \phi(a_2) \wedge ... \wedge s = \tau_A \sqcup s_1 \sqcup s_2 \sqcup ...),
\end{aligned}
$$

and

$$s \vDash^- \forall x\phi(x) \iff (\exists s' \in E \wedge \exists A \subseteq D \exists a \in A)(s' \vDash^- \phi(a) \wedge s = s' \sqcup \tau_A).$$

The above by no means constitutes a conclusive discussion of all the issues surrounding an extension of the framework to a first order semantics, but suffices for the current purposes. We have a semantics for a propositional language, and an idea of how the semantic clauses of a first order language would look like, so we can finally turn to examine how finean truth-maker semantics fair against our hyperintensional data, and, spoiling the surprise, what, if anything, can be done to improve their performance.

5.3 Conclusion

5.3.1 The Data

Let me start by examining pairs of intuitively distinct logically complex necessary truths that are conflated within the SPW framework, as for example:

$$\text{It is raining or it is not raining} \tag{5.1}$$
$$\text{It is windy or it is not windy,} \tag{5.2}$$

translated into our propositional language $\mathcal{L}$ as $R \vee \neg R$ and $W \vee \neg W$ respectively, in the obvious way. To show that finean semantics do not conflate the propositions expressed by (5.1) and (5.2) it suffices to show that there is a truthmaker for the former that is not a truthmaker for the latter, since the proposition expressed by a sentence is the pair of its set of verifiers and its set of falsifiers. According to our first semantic clause for disjunction, a truthmaker for the proposition expressed by (5.1) is any state that verifies that it is raining or that verifies that it is not raining. Let r be the state that it is raining, then r is a truthmaker for $R \vee \neg R$. However, r is not a truthmaker for $W \vee \neg W$, for r is neither a truthmaker for W nor a truthmaker for $\neg W$ since the state that it is raining is completely irrelevant to the truth of the statement that it is windy, and just as irrelevant to the truth of the statement that it is not windy. Therefore, since the propositions have distinct sets of verifiers, they are distinct propositions.

A similar argument convinces one that Fine can equally as well distinguish between intuitively distinct logically complex necessary falsehoods. Take for example the following pair:

$$\text{It is raining and it is not raining} \tag{5.3}$$
$$\text{It is windy and it is not windy,} \tag{5.4}$$

translated into our propositional language $\mathcal{L}$ as $R \wedge \neg R$ and $W \wedge \neg W$ respectively, in the obvious way. Again, it suffices to find a truthmaker for one that is not a truthmaker for the other. Let r be the state that it is raining and r' the state that it is not raining. Then the impossible state $r \sqcup r'$ is a truthmaker for $R \wedge \neg R$. However, it is not a truthmaker for $W \wedge \neg W$, since neither r nor r' is a truthmaker for W, and neither r nor r' is a truthmaker for $\neg W$. Since their sets of verifiers are distinct, the propositions expressed by (5.3) and (5.4) are distinct.

Let us now take an example of a pair of logically atomic intuitively distinct necessary truths such as $2+2=4$ and $3+3=6$. These truths are distinguished within the current framework simply because the state that $2+2=4$ is irrelevant to the truth of $3+3=6$, and hence we have a verifier for the one that is not a verifier for the other – just as was essentially the case for all previous examples. This might, however, feel a bit cheap: what we essentially have here is Fine saying that the propositions expressed by $2+2=4$ and $3+3=6$ are distinct because the state that $2+2=4$ is not the same as the state that $3+3-6$ – an explanation that might strike one as being of little explanatory value. Nevertheless, it is a coherent explanation that is formally substantiated, and perhaps the reason why it might strike one as unsatisfactory is that the possible worlds tradition has made some dubious of entities as partial as these verifying states.

But this is as far as I can go without discussing an issue that troubles me, one that becomes clear when using the framework to give a semantics for a first order language. Recall that when discussing the extension of the semantics so as to cover a first order language, the only issues we touched upon concerned adding semantic clauses and the like: there was no mention of work that needs to be done on the underlying ontological level a.k.a with respect to the state space. And this was not my call, Fine nowhere hints at issues on this level, so I assume that the state space the semantic clauses for a first order language mention is simply the extended state space extensively discussed with respect to a semantics for a propositional language. But, on this assumption, there seems to be an issue, one that comes out nicely when considering logically atomic intuitively distinct necessary falsehoods. Take the following pair:

$$\text{Amy squared the circle} \tag{5.5}$$
$$\text{Bob squared the circle,} \tag{5.6}$$

translated into $\mathcal{L}^+$ as Sa and Sb respectively, in the obvious way. The state that Amy squared the circle, as well as the state that Bob squared the circle are, to the best of my understanding, inconsistent or else impossible states; surely, a state describing a mathematical impossibility is not a possible one! Therefore, if we label the states s, s' we have that $s, s' \notin \mathcal{P}$. What is more, it is certain to my mind that they are 'simple' states, whereby a simple state I mean a state that cannot be decomposed into the fusion of other states that are proper parts of it, for what would a proper part of the state that Amy squared the circle be? Now, the passing from our state space to our extended state space was made by 'filling in the gaps', or, in other words, Fine recognises all and only those impossible states that are 'virtual fusions' of possible states. The crux of the problem is that the states s, s' do not correspond to fusions of possible states for they do not seem to have proper parts, let alone proper parts that are possible states. Therefore, states s, s' do not exist not even in our extended state space, and I fail to see what other states could possible verify *Sa* and *Sb*.[7]

Thus, the set of verifier of *Sa* is the same as the set of verifiers of *Sb* and equals the null set. This is already problematic, for Fine seems to admit too few impossibilities – we can generalise and claim that Fine fails to admit any 'atomic' or 'simple' impossibility whatsoever, he only admits impossibilities based on logical inconsistencies. However, we need to turn to falsifiers in order to conclude whether the propositions expressed by *Sa* and *Sb* are conflated. What could be a falsifier for (5.5)? Perhaps the fusion of the states that Amy did this and Amy did that and ... and that's all that Amy did. Note, here we need a restricted totality claim – the restriction being to the property of *being something that Amy did*, which can be quite tricky to construct and justify. Nevertheless, indeed, this is no falsifier of (5.6). Content being bilateral potentially saves Fine from conflating these intuitively distinct impossibilities. Yet Fine's account admits fewer impossibilities than there intuitively are.

Just as for Yablo, I do not go into the hyperintensional data pertaining to embeddings in epistemic, opaque or non-opaque contexts. I merely note that pairs of propositions we usually embed in such contexts to get variants of Frege's puzzle for example, or other results where truth values predicted by a framework misalign with intuitive truth judgments, are all pairs that Fine one way or another distinguishes between whose members to begin with. Therefore, even if a purely relational semantics for such operators is developed, the entities/ propositions agents are going to be related to being distinct, we can be confident that truth-value intuitions about the embeddings can be preserved.

One final issue I want to touch upon is that the current 'liberal' and completely unconstrained definition of propositionhood makes hyperintensional distinctions where we would like none – it delivers, so too speak, too fine a grain. Take for example R "It is raining" and $R \wedge R$ "It is raining and it is raining". The sole verifier of R may plausibly be taken to be the state r that it is raining. But it that case, $r \sqcup r$ is the sole verifier of $R \wedge R$, thus the propositions expressed by the two sentences above differ – a rather unintuitive result. However, in Fine's setting it is not that hard to constrain the finess of grain from above, contrary to what was the case with proponents of structured propositions. For example, the above result can be blocked by requiring that any set of verifiers that figures in the first place of a bilateral content be a set that satisfies 'Closure'. I believe it is worth a note, but this is not an issue I will be focusing on: Fine seems to be well aware of it, and many a suggestions have already been made.[8] Calibrating the grains by debating exactly what conditions to impose on sets of verifiers/ falsifiers so that they pair into contents is something that needs to be done, but seems to be something that can be done – that it hasn't yet could be a complaint but not a serious one.

5.3.2 The Verdict

Fine's account fairs well with the data and does not seem subject to crippling issues of a general nature. If anything, the account is elegant, formally substantiated and much suited to the work, especially if contrasted with the Yablo-Jago account. States being partial lend themselves to the accommodation of aboutness, and the passing from the possible to the impossible ones seems the contrary of *ad hoc* – rather, we complete the space.

[7]Note that this goes against 'Verifiability', a principle that I nevertheless did not pick to characterise propositions.
[8]See p.10 of Fine 2016c.

The greatest issue I take with Fine's account is that it delivers too few impossibilities; not too few so to be strongly hyperintensional – for bilateral content does the trick – but simply, and more fundamentally and worrisomely so, fewer impossibilities than we intuitively think there are. This issue reproduces of course on the level of content. For example, if Fine wants a quantificational rendering of modal truths, then I see no way for "It is impossible that Amy squared the circle" to be true. There exists no state that verifies that Amy squared the circle – no impossible one, surprisingly – and hence no state-based entity such that it supports the truth of the claim. But this is not an issue best discussed on the level of content; I insist that the problem is readily and best identified on the ontological level: there are too few impossible states.

Fine is very clear on how to arrive to the totality of impossible states: non-principal ideals of possible states correspond to impossible states whose existence is then postulated. The problem then is with the extension. But if we abandon these means of extending the state space, Fine's account plausibly collapses into something very akin to the impossible world semantics framework. I take no issue with impossible states delivered *à la* Fine; I simply claim we need more. So let us postulate that, other than impossible states corresponding to 'virtual fusions', $E \setminus \mathcal{P}$ contains, for each necessary falsehood we intuitively want that is not otherwise verifiable, at least one simple impossible state verifying it. This readily solves the ontological issue, but destroys a motivation of the framework that Fine counts on.

More precisely, Fine takes issue with impossible world semantics Priest-style because, although they come with a more refined conception of entailment, and thus a more refined conception of mutual entailment and propositions, "[i]t is not likely that we want to jettison all non-trivial entailments or all non-trivial identities between propositions (Fine 2016b, p.2)" – as we have seen that that semantics indeed does. This is very much the same issue Jago takes with Priest-style extended possible world semantics, discussed under the guise of the problem of bounded rationality. Fine's argument goes that the only remedy for extended possible world semantics is to restrict the range of impossible worlds, but that the only means of doing so in an way that is not completely *ad hoc* is to somehow have the possible serve as a guide to the impossible. Leave it aside for now that in Jago we have found a different strategy that is not so obviously *ad hoc* as the selected worlds approaches that Fine considers – we have in Jago not a selection of worlds based on what rules of inference the worlds are closed under, but rather a selection of worlds based on where they fall on an ordering imposed on their totality by means of some proof system or another, the choice of proof system being detestably *ad hoc*. Let us focus instead on the following: the strategy Fine uses to get to his totality of impossible worlds is one that uses the possible as a guide to the impossible. However, by imposing the addition of 'simple' impossible states to his totality of states, the possible is no longer a guide to the impossible. And indeed, perhaps Fine is right in his diagnosis as to why impossible world semantics suffer from too weak an entailment, for as soon as we add impossible states without being guided by the possible, any entailment Fine's semantics supports is supported by extended possible world semantics *à la* Priest.

In other words, what we have is that, where $|X|_F$ is the proposition that X *qua* Fine, $|X|_P$ the proposition that X *qua* Priest, $\vDash_F, \vDash_P$ the truthmaking relation *qua* Fine and Priest respectively:[9]

$$\text{if } |P|_F \text{ Fine-entails } |Q|_F, \text{ then } |P|_P \text{ Priest-entails } |Q|_P. \tag{5.7}$$

To remind the reader, $|P|_P$ Priest-entails $|Q|_P$ if and only if $|P|_P \subseteq |Q|_P$, and I take it that entailment is defined set theoretically for Fine too: $|P|_F$ Fine-entails $|Q|_F$ if and only if $|P|_F \subseteq$

[9] I have switched here to a unilateral account of content on Fine's side, whereby a proposition is identified with its set of verifiers, and dually I identify contents with positive contents a.k.a sets of worlds that make true a given content on Priest's side. This is for ease of exposition, but the argument that follows could be duplicated for sets of falsifiers and negative contents.

$|Q|_F$.[10] Before proceeding to show that (5.7) is the case, let me make clear what it proves for Fine's account: that it preserves at most as few inferences as Priest's does, the latter having been argued already too few by Fine himself. The reason this problem arises when we add 'simple' impossible worlds to the mix and not with Fine's original extension is that we now have a bijective mapping between Priest's totality of possible worlds W and sets of states from Fine's totality of states $\mathcal{S}$:

$$\begin{aligned} &f : W \to \mathcal{P}(\mathcal{S}) \\ &\text{such that } w \vDash_P A \iff (\exists s \in f(w))(s \vDash_F A) \end{aligned} \tag{5.8}$$

The proof is straightforward: assume that $|P|_F \subseteq |Q|_F$, and take an arbitrary $w \in W$ such that $w \vDash_P P$. Now let $f(w) = S \subseteq \mathcal{S}$. From our assumption and definition of f, it follows that $(\exists s \in S)(s \vDash_F P)$, which, given $|P|_F \subseteq |Q|_F$, gives us that $s \vDash_F Q$. But then, from the definition of f we have that $w \vDash_P Q$ as required.

Let us recap. It has been argued that Fine's totality of states includes too few impossible states. The sort of impossible states it does not include while it should have been postulated into existence and added to the state space, in the most natural way the account can be regimented to dispense with the problem. But the strategy has given rise to an important issue: just as impossible world semantics *à la* Priest destroy too many an inferences, so do the current state semantics. Fine himself has argued against impossible world semantics using this very consequence of the framework, so I know not what to conclude. We can either leave it as we found it and say that Fine's account suffers a great ontological deficit that results in him not being able to represent many an impossible scenarios that we intuitively want, or we can say that his account, regimented in the most natural way, falls pray to the very same criticism he deems damning when raised against impossible world semantics.

So, for the big question, how does Fine's account compare with the Yablo-Jago one? Both deal equally well with all the data, so on the hyperintensionality front they score the same high score. However, Fine's account suffers from an obvious drawback – it has not found a way to accommodate for the impossible (not all of it) unless it compromises with rendering too few inferences valid. Let us not forget, however, that the original framework Jago's work is a contribution towards suffers from the exact same condition. Jago's is a highly sophisticated and ideologically loaded solution to the problem (if not covertly *ad hoc*) that needs to take on even more ideology – that of Yablo's – to deliver the results that Fine's account does, simply in virtue of operating on the basis of partial entities. Add to that that Yablo's account of truthmaking is more problematic than Fine's (both are instrumentalist, but at least Fine's does not purport that essentially unanalysed entities play a foundational role in semantics), and the scale is hard to read. I for one, and this is admittedly a prediction more than it is a well-founded conclusion, see hope in Fine-style state-based truthmaker semantics: the theory has not been around as long and has not been worked on as much as the intensional framework Yablo's and Jago's works are based on yet it delivers the same finess of grain, and it seems better suited from the very onset to make hyperintensional distinctions (recall that even the propositions construed out of the (non-extended) state space cut more finely than SPW propositions) due to the partial nature of its base entities.

[10] See Fine 2016c §6 for a more detailed discussion of consequence; I have here simplified the case he makes for purposes of exposition.

Chapter 6

Conclusion

I conclude in a recap. In the first chapter of the current thesis I argued that a satisfactory account of meaning must be hyperintensional, and concluded on a list of pairs of sentences I have been calling 'the data' that help keep tabs on and demarcate certain hyperintensional distinctions I have argued a good theory of content should be able to draw. In the main body of the thesis, with the 'data' as a guide, I examine four different accounts of meaning which have made claims to being hyperintensional, my aim being to discover which, if any, is an account of meaning that is theoretically virtuous and, more to the focus, grinds propositions just finely enough.

In chapter 1, I examined some members of a family of accounts of meaning – those that take meanings to be structured entities. The investigation led me to concluded that the most satisfying variant of structuralism – Russellianism *à la* Soames and Salmon – fairs well with the data, but refuses to recognise the need to distinguish between intensionally equivalent propositions expressed by sentences that differ only in so far as they employ distinct coreferential terms. Adopting this stance is asking us to explain away some rather strong Fregean intuitions, and the explanations set forth by Soames and Salmon are not entirely uncontroversial. Moreover, the account seems to cut too finely at points. Add to that the most promising attempt to address the question of the unity of propositions, King's from §2.2.2, is partly self-defeating – in that it takes avoiding Benacerraf-style worries to be a main motivation for adopting the account, yet is faced with 'revenge' Benacerraf concerns – and delivers many a counterintuitive consequences stemming from the assumption embedded in the account that propositions are ontologically dependent on language.

Structuralism is thus found wanting, and from the second chapter onwards I turn to non-structured accounts of meaning. Starting with extended possible world semantics in Chapter 2, I conclude that the most satisfying account is that of Jago's, but that it seems to be committed to conflating certain necessary truths and falsehoods, violating intuitions of aboutness. Jago's account being generally theoretically virtuous – other than perhaps slightly *ad hoc* – and fairing well with the rest of the data, was put aside for the time. In the next chapter, I examined an account of meaning based on Yablo's work in Yablo 2014a which albeit an intensional, non-structured account of meaning manages to model aboutness and capture many a related intuitions. It is nevertheless a worlds based semantics that does not allow for impossible worlds, and it was thus found to conflate necessarily equivalent and logically atomic statements. However, combining Yablo's and Jago's work into a single semantics, we get propositions that seem to have just the right granularity, in a move that dispenses with Jago's issues, dispenses with Yablo's issues and creates no new ones but for the fact that one might start wondering whether we have too much ideology.

In Chapter 4 I examined an alternative truthmaker semantics based on a state space developed by Kit Fine in a series of articles. Fine's account fairs well with all the data, but faces a problem reminiscent of that Jago raises against Priest: either we have too few impossible states, or we have a semantics that destroys too many an inferences. The Jago-Yablo account suffers lesser drawbacks and deals equally well with the data, yet I am disinclined to conclude it is the better one for the following reasons: a Finean account of content having partial entities such as states on its basis seems better suited for delivering hyperintensionality in a less contrived and ideologically loaded way, as well as more deserving of the title of a truthmaker semantics. Moreover, it already cuts as fine – albeit still in its infancy compared to possible world based semantics such as Jago's and Yablo's – thus inspiring justified hope.

Bibliography

Armstrong, David Malet (2004). *Truth and Truthmakers*. New York: Cambridge University Press.

Barwise, Jon and John Perry (1983). *Situations and Attitudes*. Cambridge, MA: MIT Press.

Bealer, George (1998). “Propositions”. In: *Mind* 107.January, pp. 1–32.

— (2004). “An Inconsistency in Direct Reference Theory”. In: *Journal of Philosophy* 101.11, pp. 574 –593.

Beaver, David I. and Geurts Bart (2014). “Presupposition”. In: *The Stanford Encyclopedia of Philosophy*.

Berto, Francesco (2013). “Impossible Worlds”. In: *The Stanford Encyclopedia of Philosophy*. URL: `http://plato.stanford.edu/archives/win2013/entries/impossible-worlds/`.

Branquinho, Joao (1990). “Are Salmon’s ‘Guises’ Disguised Fregean Senses?” In: *Analysis* 50, pp. 19–24.

Chalmers, David J (2011). “Propositions and Attitude Ascriptions: A Fregean Account”. In: *Noûs* 45.4, pp. 595 –639.

Ciardelli, Ivano, Jeroen Groenendijk, and Floris Roelofsen (2013). “Inquisitive semantics: A new notion of meaning”. In: *Linguistics and Language Compass* 7.9, pp. 459–476. ISSN: 1749818X.

Crimmins, Mark and John Perry (1989). “The Prince and the Phone Booth : Reporting Puzzling Beliefs”. In: *The Journal of Philosophy* 86.12, pp. 685–711.

Deutsch, Harry (2008). “Jeffrey C.King, The Nature and Structure of Content (2007), Reviewed by Harry Deutch”. In: *Notre Dame Philosophical Reviews*.

Eddon, Maya (2011). “Intrinsicality and Hyperintensionality”. In: *Philosophy and Phenomenological Research* 82, pp. 314–336.

Fine, Kit (1994). “Essence and Modality”. In: *Philosophical Perspectives* 8, pp. 1–16.

— (2012). “Guide to Ground”. In: *Metaphysical Grounding*. Ed. by Fabrice Correia and Bill Schnieder. Cambridge, MA: Cambridge University Press, pp. 37–80.

— (2014). “Truth-Maker Semantics for Intuitionistic Logic”. In: *Journal of Philosophical Logic* 43.2-3, pp. 549–577.

— (2015). “A Theory of Truth-Conditional Content I: Conjunction, Disjunction and Negation”. URL: `https://www.academia.edu/9186494/Truthconditional{\_}Content{\_}-{\_}Part{\_}I`.

— (2016a). “Angellic Content”. In: *Journal of Philosophical Logic* 45, pp. 199–226.

— (2016b). “Constructing the Impossible”. URL: `https://www.academia.edu/11339241/Constructing{\_}the{\_}Impossible`.

— (2016c). “Truthmaker Semantics”. URL: `https://www.academia.edu/8980719/Truthmaker{\_}Semantics`.

— (2016d). “Yablo on Subject-Matter”. URL: `https://www.academia.edu/15555407/Review{\_}of{\_}Steve{\_}Yablos{\_}Aboutness`.

Forbes, Graeme (1987). “FREGE’S PUZZLE. By Nathan Salmon. Cambridge, Mass,:The MIT Press, 1983. Review by G.Forbes”. In: *The Philosophical Review* 96.3, pp. 455–458.

Frege, Gottlob (1892). “On Sense and Meaning”. In: *Frege: Collected Works*. Ed. by B. McGuinness. Oxford: Basil Blackwell, pp. 157–177.
Hale, Bob (2007). “Critical Studies / Book Reviews: Into The Abyss”. In: *Philosophia Mathematica* 15.3, pp. 94–134.
Hanks, Peter (2009). “Recent Work on Propositions.” In: *Philosophy Compass* 4.3, pp. 469–486.
Hazen, Allen P. and Lloyd Humberstone (2004). “Similarity Relations and the Preservation of Solidity”. In: *Journal of Logic, Language and Information* 13.1, pp. 25–46.
Hofweber, Thomas (2006). “Inexpressible Properties and Propositions”. In: *Oxford Studies in Metaphysics*. Ed. by Dean Zimmerman. Oxford: Oxford University Press, pp. 155–2006.
Hughes, G.E. and M.J. Cresswell (1996). *A New Introduction to Modal Logic*. New York: Routledge.
Jago, Mark (2014). *The Impossible: An Essay on Hyperintensionality*. Oxford, UK: Oxford University Press.
Jenkins, Carrie and Daniel Nolan (2012). “Disposition Impossible”. In: *Noûs* 46.4, pp. 732–753.
Jubien, Michael (2001). “Propositions and the Objects of Thought”. In: *Philosophical Studies* 104.1, pp. 47–62.
Kaplan, David (1978). “Dthat”. In: *Syntax and Semantics*. Ed. by P Cole. New York: Academic Press, pp. 221–253.
King, Jeffrey (1995). “Structured propositions and complex predicates”. In: *Nous* 29.4, pp. 516–535.
— (2007). *The Nature and Structure of Content*. Oxford, UK: Oxford University Press.
— (2013). “Propositional unity: What’s the problem, who has it and who solves it?” In: *Philosophical Studies* 165.1, pp. 71–93.
Kripke, Saul (1980). *Naming and Necessity*. Oxford, UK: Basil Blackwell.
Lewis, David (1973). *Counterfactuals*. Oxford: Blackwell Publishers.
— (1980). “Index, Context, and Content”. In: *Philosophy and Grammar*. Ed. by Stig Kanger and Sven Ohman. Dodrecht: Reidel, pp. 79–100.
— (1988). “Statements Partly About Observation”. In: *Philosophical Papers* XVII.I, pp. 1–31.
— (2004). “Letters to Priest and Beall”. In: *The Law of Non-Contradiction: New Philosophical Essays*. Ed. by B Armour-Garb, J Beall, and Graham Priest. Oxford: Oxford University Press, pp. 176–177.
MacBride, Fraser (2016). “Truthmakers”. In: *The Stanford Encyclopedia of Philosophy*.
McGrath, Matthew (2014). “Propositions”. In: *The Stanford Encyclopedia of Philosophy*.
Nolan, Daniel (2013). “Hyperintensional metaphysics”. In: *Philosophical Studies* 171.1, pp. 149–160.
Pastan, Stephen (2016). “Another Approach to Truthmaker Semantics”. PhD thesis. ILLC, University of Amsterdam.
Priest, Graham (1997). “Sylvan’s Box: A Short Story and Ten Morals”. In: *Notre Dame Journal or Formal Logic* 38, pp. 573–581.
— (2005). *Towards Non-Being: The Logic and Metaphysics of Intentionality*. Oxford: Oxford University Press.
— (2014). *One: Being an Investigation into the Unity of Reality and of its Parts, including the Singular Object which is Nothingness*. Oxford: Oxford University Press.
Ripley, David (2012). “Structures and circumstances: Two ways to fine-grain propositions”. In: *Synthese* 189.1, pp. 97–118.
Russell, Bertrand (1918). “The Philosophy of Logical Atomism”. In: *Logic and Knowledge*. Ed. by R. C. Marsh. London: Allen and Unwin, pp. 177–281.
— (1972). *The Philosophy of Logical Atomism*. Routledge. Routledge.
Salmon, Nathan (1986). *Frege’s Puzzle*. Cambridge, MA: MIT.

Schaffer, Johnathan (2009). “On What Grounds What”. In: *Metametaphysics*. Ed. by David John Chalmers, David Manley, and Robert Wasserman. New York: Oxford University Press, pp. 347–383.
Schnieder, Benjamin (2011). “A logic for 'Because'”. In: *The Review of Symbolic Logic* 4.3, pp. 445–465.
Soames, Scott (1985). “Lost Inoccence”. In: *Linguistics and Philosophy* 8.1, pp. 59–71.
— (1987). “Direct Reference, Propositional Attitudes, and Semantic Content”. In: *Philosophical Topics* 15, pp. 47–87.
— (2002). *Beyond Rigidity: The Unfinished Semantic Agenda of Naming and Necessity*. Vol. 1. New York: Oxford University Press.
— (2008). “Why Propositions Cannot be Sets of Truth- Supporting Circumstances”. In: *Journal of Philosophical Logic* 37, pp. 267–276.
Stalnaker, Robert (1968). “A Theory of Conditionals”. In: *American Philosophical Quarterly*, pp. 98–112.
— (1976). “Possible Worlds”. In: *Noûs* 10.1, pp. 65–75.
— (1987). “Semantics for Belief”. In: *Philosophical Topics* 15.1, pp. 177–190.
Strawson, Peter F. (1950). “On Referring”. In: *Mind* 59, pp. 320–344.
Van Frassen, Bas (1969). “Facts and Tautological Entailments”. In: *The Journal of Philosophy* 66.15, pp. 477–487.
Yablo, Stephen (2014a). *Aboutness*. Princeton University Press.
— (2014b). *Aboutness Theory*. URL: `www.mit.edu/{\~{}}yablo/home/Papers{\_}files/aboutnesstheory.pdf`.
— (2016). “Ifs, Ands, and Buts: An Incremental Truthmaker Semantics for Conditionals”. In: *Analytic Philosophy*.
Yagisawa, T (1988). “Beyond Possible Worlds”. In: *Philosophical Studies* 53, pp. 175–204.