UNIVERSIDADE DE LISBOA

FACULDADE DE LETRAS

On the Topic of Impossibility

a question-sensitive impossible worlds approach to logical omniscience

Francisca Silva

Dissertação orientada pelo Prof. Doutor Ricardo Santos, especialmente elaborada para a obtenção do grau de Mestre em Filosofia.

2021

Acknowledgements

First and foremost, I would like to thank Professor Ricardo Santos for supervising this dissertation. His detailed comments helped improve at various points and to a large degree the arguments I've developed and the overall discussion. Likewise, his fair criticism often made me consider certain options more carefully as well as reconsider certain positions I first thought plausible. I am very grateful for Professor Ricardo's continued trust in my philosophical work, as well as for the candid but kind and motivating way he has commented on several drafts of this dissertation.

Mark Jago made numerous invaluable comments, questions and interesting suggestions for how to take this project further in a number of ways I hadn't begin to consider. I am both flattered and immensely thankful for the way he closely read this dissertation, as well as for the resulting rich and illuminating discussion I was able to enjoy during the *viva voce* having him as the examiner.

I was fortunate to have the opportunity to present some of the views here expounded in the LanCog's Seminar Series in Analytic Philosophy. I would like to thank everyone in the audience for their questions and suggestions, which were very helpful and illuminating. Ever since I became a student member of LanCog roughly two years ago, I've been welcomed to a very rich and stimulating environment, having the opportunity to be in contact with several ideas and contemporary debates in various areas of Philosophy both through the aforementioned seminar series, workshops, as well as the reading groups organized by some of its members. I am very grateful for continuing to be a part of LanCog and for how it has helped me improve over the last two years.

I was first introduced to the issues that the dissertation addresses, to a large extent, upon attending a reading group in Epistemology, organized by Bruno Jacinto. I am very thankful to Bruno for having introduced me to these questions, which have occupied me since. I would also like to especially thank Bruno for the numerous occasions in which he was available to discuss some of the perspectives here considered and answer my questions about them. These discussions have for certain greatly improved the dissertation, but they've also, I like to believe, helped improve my dialectical skills.

I would like to thank Franz Berto both for the "free supervision session" as well as for being available to answer some questions in a very approachable way. His suggestions of further papers/books/dissertations to read pointed me in very helpful directions, some of which I regret I have only been able to explore in a very limited way at this stage.

I would also like to thank everyone from the Argument Clinic student group who attended the discussions on *On the Plurality of Worlds*, and especially to those who have accepted to present their own work; these informal sessions were of great help in early stages of development of this dissertation. I would like to thank Domingos Faria for encouraging these activities as well as for all the support, as well as José Mestre, Bruno Jacinto and Diogo Santos, for trusting me to help out with the group. Here I would like to again particularly thank Bruno

Jacinto for making himself available to give an informal mini course on possible world semantics and second order logic, tools which were essential for a better understanding of the frameworks here analyzed and compared.

My friends from the Minorities and Philosophy chapter at the Faculty of Letters of the University of Lisbon, Ana Logrado, João Amorety, João Patrocínio and Sofia Estudante have often been an emotional pillar during the last months whilst at the same time often reminding me of various important issues that need to be actively dealt with in the actual world, helping me not get lost in the abstract space of impossible worlds. For this, I am very thankful.

Allie has been an invaluable friend, without whose support I would never have been able to complete this dissertation, or at least certainly not in a way I would have been satisfied with. I am very thankful to her for her friendship and care. What, if anything, has been accomplished in this dissertation is as much my achievement as it is hers.

I would like to extend a heartfelt thank you to my parents, Vladimiro and Anita, for always believing in me, for their support and for their patience and understanding in dealing with some of the misadventures of having a Philosophy student for a daughter.

Finally, I would like to thank the Centre of Philosophy of the University of Lisbon (CFUL) for a scholarship which I have been fortunate to have been granted for the last year. The financial support provided by the CFUL helped make this dissertation possible in a way that cannot be overstated.

To Inês,
In fulfilment of an old promise.

Resumo

Desde que Hintikka (1962) propôs uma lógica epistémica modal com base na semântica de mundos possíveis que rapidamente se colocou o problema da omnisciência lógica, isto é, de que a lógica em questão implica que os agentes sabem tudo o que se segue logicamente do que sabem. O próprio Hintikka (1975) tentou resolver o problema introduzindo "mundos possíveis impossíveis". Desde então, os mundos impossíveis têm sido aplicados no tratamento de várias outras questões filosóficas. Berto e Jago (2019) desenvolvem e exploram várias delas.

A presente dissertação começa com uma avaliação em detalhe de soluções para o problema da omnisciência lógica que aceitam mundos impossíveis. De forma a melhor considerar esta perspetiva, questões sobre a caracterização da natureza dos mundos impossíveis e de como representam são consideradas.

Por outro lado, filósofos como Yalcin (2018) propõem dar resposta ao problema da omnisciência lógica sem acrescentar mundos impossíveis, mas sim tendo por base uma extensão da noção do conteúdo de frases e estados mentais. De acordo com esta segunda família de perspetivas, o conteúdo de uma frase não é dado simplesmente em termos de condições de verdade, mas também em termos daquilo sobre aquilo que a frase versa. O conhecimento de agentes seria fechado sob implicação lógica que não adiciona nenhum novo assunto ao das proposições que o agente sabe, mas não sob implicação lógica *simpliciter*.

Esta segunda família de perspetivas será igualmente considerada, começando pela questão de o que são assuntos, e terminando por considerar se as várias perspetivas disponíveis conseguem dar conta de todas as diferenças entre conteúdos face aos quais agentes podem ter atitudes proposicionais distintas.

Finalmente, será desenvolvida uma solução para o problema da omnisciência lógica que aceita tanto mundos impossíveis, como uma relativização a questões ou assuntos.

Palavras-chave: mundos impossíveis; assuntos; omnisciência lógica; lógica epistémica.

Abstract

Ever since Hintikka (1962) proposed an epistemic logic based on possible worlds semantics, the problem of logical omniscience, that is, that the logic proposed by Hintikka would have as a consequence that agents know all the logical consequences of what they know, has been posed as a challenge. Hintikka (1975) himself tried to meet the challenge by introducing "impossible possible worlds". Since then, impossible worlds have been applied to the treatment of various philosophical questions. Berto and Jago (2019) develop and explore several of them.

The present dissertation starts by considering in detail solutions for it that accept impossible worlds. In order to more fully consider this family of perspectives, questions regarding the nature of impossible worlds and how they represent are discussed.

On the other hand, philosophers like Yalcin (2018) propose to give a solution to the problem of logical omniscience without adding impossible worlds to a standard possible worlds framework, but rather accepting an extension of the notion of the content of statements and mental states. According to this second family of perspectives, the content of a statement is not simply given in terms of truth-conditions, but also in terms of what the statement is about. Agents' knowledge would be closed entailment that does not add any new subject matter to the subject matter of what the agent knows, not under logical consequence *simpliciter*.

This second family of perspectives will also be considered, starting from the question of what subject matters and including others, such as whether various perspectives on offer are able to account for all the distinctions between contents to which agents might have different propositional attitudes.

Finally, a solution for the problem of logical omniscience that accepts both impossible worlds and subject matters will be developed.

Keywords: impossible worlds; logical omniscience; subject matters; epistemic logic.

Resumo alargado

Desde que Hintikka (1962) propôs uma lógica epistémica modal com base na semântica de mundos possíveis que rapidamente se colocou o problema da omnisciência lógica, isto é, de que a lógica em questão implica que os agentes sabem tudo o que se segue logicamente do que sabem. O próprio Hintikka (1975) tentou resolver o problema introduzindo “mundos possíveis impossíveis”. Desde então, os mundos impossíveis têm sido aplicados no tratamento de várias outras questões filosóficas. Mais recentemente Berto e Jago (2019) desenvolveram e exploraram várias delas, apresentando igualmente uma perspetiva aprofundada a respeito da natureza dos mundos impossíveis.

A presente dissertação começa com uma avaliação em detalhe de soluções para o problema da omnisciência lógica que aceitam a existência de mundos impossíveis. De forma a melhor considerar esta perspetiva, questões sobre a caracterização da natureza dos mundos possíveis e de como representam começam por ser consideradas, assim como quais perspetivas são satisfatórias por si mesmas, e quais poderão ser expandidas de forma a acomodar igualmente mundos impossíveis.

Começando por notar, face à forma como mundos representam, a distinção entre mundos genuínos e mundos ersatz, concluir-se-á rapidamente que os mundos impossíveis não poderão ser mundos genuínos. Passamos então a considerar várias perspetivas ersatzistas a respeito de como mundos representam. A primeira aqui considerada, o ersatzismo mágico, rejeita a necessidade de dar conta da forma como mundos representam, tratando como primitivo o facto de que certos mundos representam aquilo que representam. Como se verá, esta posição sofre de uma objeção decisiva por parte de Lewis (1986), de acordo com a qual apenas por magia é possível compreender as noções primitivas da mesma.

Após uma defesa das críticas de Lewis a esta perspetiva, passa-se então a considerar a família de perspetivas ersatzistas construtivistas, de acordo com as quais os mundos possíveis são construções abstratas. Como se verá, a objeção que Lewis apresenta para o ersatzismo mágico igualmente se aplica às perspetivas construtivistas em que os constituintes últimos dos mundos possíveis são eles mesmos abstratos e simples. Igualmente, outras considerações levam-nos a rejeitar posições de acordo com as quais mundos possíveis são formados por proposições e estados de coisas. Conclui-se assim que a melhor perspetiva a respeito da natureza dos mundos possíveis é o ersatzismo linguístico, isto é, a perspetiva de acordo com a qual os mundos possíveis são conjuntos maximamente consistentes de frases de uma linguagem lagadoniana, uma linguagem em que os próprios objetos, propriedades e relações correspondem respetivamente aos seus termos singulares e gerais.

Após se considerarem as várias formas de ersatzismo, uma forte objeção a perspetivas ersatzistas em geral, o problema dos objetos e propriedades alienígenas, é discutida. Em resposta ao mesmo, é aceite uma estratégia descritivista, em que objetos/propriedades que não existem mas poderiam existir são representados por descrições definidas da linguagem lagadoniana. Estas descrições definidas, irá propor-se, contêm uma referência ao mundo possível em que seriam satisfeitas, sendo que os objetos e propriedades são identificados por serem aqueles que satisfariam uma determinada descrição definida caso um determinado mundo possível fosse o mundo atual. No entanto, dado que estas descrições definidas são expressões de uma linguagem

lagadoniana, a referência ao mundo possível em que são satisfeitas leva a preocupações de circularidade, pois o mundo possível em questão seria ele próprio um conjunto de frases lagadonianas algumas das quais incluirão a descrição definida em questão. De forma a dar conta desta objeção, e a melhor compreender a natureza dos mundos possíveis, são então discutidas questões relativas à natureza e mereologia dos conjuntos, assim como a possibilidade de adotar uma mereologia, motivada independentemente, que rejeita que a relação de parte que seja anti-simétrica. Da mesma resulta a visão de conjuntos como fusões de fusões de Fine (definidas em termos de certos objetos instanciarem certas propriedades e relações), que permite tanto evitar a objeção de Lewis considerada anteriormente, como oferecer justificações para se aceitarem certas restrições à formação de conjuntos, que bloqueam conjuntos como o ínfame conjuntos de todos os conjuntos que são membros de si próprios.

Com uma visão robusta de ersatzismo linguístico para o caso dos mundos possíveis, passamos a averiguar como esta perspetiva poderá ser extendida de forma a acomodar mundos impossíveis. Destas considerações extrai-se uma visão ontologicamente inocente dos mundos impossíveis em que aceitar tanto mundos impossíveis como mundos possíveis é simplesmente aceitar certas construções abstratas que têm objetos atuais como base. Na perspetiva resultante, os mundos possíveis correspondem a conjuntos maximamente consistentes de frases da linguagem lagadoniana, e os mundos impossíveis a todos os restantes. Assim, neste enquadramento aceitar a existência de mundos impossíveis não representa um maior compromisso ontológico face a aceitar mundos possíveis: se se aceitam conjuntos maximamente consistentes de frases, porque não aceitar todos os restantes?

Na posse de uma conceção de mundos como conjuntos de frases de uma linguagem lagadoniana, passou-se a considerar como uma perspetiva que aceita mundos impossíveis poderá dar resposta ao problema da omnisciência lógica. Comparando as perspetivas de Bjerring e Skipper (2018) e Berto e Jago (2019), chegou-se a uma visão que incorpora elementos de ambas, favorecendo uma solução em que o que agentes conhecem depende de que questões consideram, e por isso que um modelo adequado do conhecimento de agentes limitados terá necessariamente uma componente dinâmica.

Uma vez que a noção de questão/assunto não havia sido esclarecida, na segunda parte da dissertação foram consideradas perspetivas sobre a natureza de assuntos e questões, assim como duas propostas de solução para o problema da omnisciência lógica, a de Yalcin (2018) e a de Hawke, Özgün e Berto (2020), que não acrescentam mundos impossíveis a um enquadramento com mundos possíveis, mas que têm por base, antes, uma extensão da noção do conteúdo de frases e estados mentais. De acordo com esta segunda família de perspetivas, o conteúdo de uma frase não é dado simplesmente em termos de condições de verdade, mas também em termos daquilo sobre aquilo que a frase versa. O conhecimento de agentes seria fechado sob implicação lógica que não adiciona nenhum novo assunto ao das proposições que o agente sabe, mas não sob implicação lógica *simpliciter*.

Da análise destas perspetivas resultou que todas elas ganham em plausibilidade com a adição de mundos impossíveis na caracterização que apresentam da noção de assuntos/questões, assim como nos seus modelos para o conhecimento para agentes limitados, uma vez que apenas com mundos impossíveis será possível distinguir entre conteúdos a respeito dos quais agentes podem ter atitudes epistémicas distintas. Chegou-se, assim, à conclusão de que a plausibilidade de soluções com base em mundos impossíveis requer a aceitação de que o que agentes sabem depende de que questões consideram, e por outro lado que soluções com base na noção de assunto requerem mundos impossíveis.

Com base na discussão precedente, é finalmente apresentada uma solução para o problema da omnisciência lógica que recorre a vários elementos das perspetivas até então consideradas. Um agente é nesta perspetiva modelado como apreendendo proposições na medida em que é capaz de considerar os factos expressos pelas frases da linguagem lagadoniana correspondentes, e como focando-se em diferentes conjuntos de mundos dependendo das questões que consideram. Estes conjuntos encontram-se por sua vez divididos numa partição em que cada uma das celas corresponde a uma forma do mundo ser relativamente ao assunto ou alternativamente a uma resposta completa à questão que o agente considera.

Table of contents

Introduction

Possible worlds semantics has been widely applied in efforts to provide a more precise understanding of the nature of propositions, as well as of various kinds of possibility and necessity. A result of this application is the influential conception of propositions as sets of possible worlds, and, in a very direct way, from this understanding the conception that propositional attitudes, such as knowledge and belief, are relations between agents and sets of possible worlds. An objection that was soon raised to this last hypothesis is that it validates various closure principles for, among others, knowledge and belief that do not seem to apply to non-ideal agents like the average adult human being. Three closure principles generally taken to be best avoided run as follows[1]:

Closure under believed conjunction

(BC) If A believes that p and A believes that q, then A believes that $p \wedge q$.

Closure under logical implication

(LI) If A believes that p, and p logically implies that q, then A believes that q.

Closure under necessary equivalence

(NE) If A believes that p, and p and q are necessarily equivalent, then A believes that q.

Even some of the most notable proponents of the sets of possible worlds conception of propositions take (BC) to be a problematic closure principle that we would do best to avoid. Stalnaker (1984) and Lewis (1982, p. 436) give examples that go against this closure principle, which come down to cases where a given cognitive agent fails to put together the information they have available. Let us take Lewis's example: while believing that Nassau Street ran roughly East-West and the nearby train station ran roughly North-South, Lewis also believed them to be roughly parallel – a contradiction. Here, while believing each of the conjuncts in the triple, Lewis

[1] Here the versions of these principles applied to belief are presented, but they apply equally well to knowledge. In fact, given that knowing p implies belief in p for any proposition p and any plausible account of knowledge, it suffices to show that the corresponding closure principles for belief yield unacceptable results to conclude that they also should not be upheld for knowledge. If a given closure principle is taken to apply to belief overall, then it is also taken to apply to those beliefs that amount to knowledge. Hence, if it is further shown that the closure principle under consideration is unacceptable as a closure principle for belief, then so is it unacceptable as a closure principle for knowledge. The point here expressed does not rest on the idea that knowledge is further analysable into notions such as belief, truth and justification. Rather, it rests only on the intuition that knowledge that p always implies belief that p. This is accepted even by knowledge-first approaches in Epistemology, as in Williamson (2000).

did not believe their conjunction. Lewis did not believe any contradiction, and yet his beliefs contradicted each other. Closure under conjunction must, then, fail, for otherwise Lewis could not come to hold beliefs in each of the conjuncts, but not in their conjunction.

While examples such as the one presented by Lewis might convince us of the unacceptability of (BC), for what has been said so far, we seem to still lack a story for how common situations that limited agents face are such that they fail to put together their beliefs into a belief in their conjunction. The diagnosis proposed by Lewis is that one might believe a number of propositions according to different fragments of one's belief systems (which are called upon at different times), and not believe that their conjunction is the case in any fragment (and hence not believe at all, since agents only believe in propositions from the standpoint of one fragment or other). In each of these fragments, we have closure of belief under conjunction, as well as closure under classical logical implication and necessary equivalence. Non-ideal agents in non-ideal contexts are computationally limited and, for instance, must perform actions within a limited time frame; and therein comes the fragmentation. According to this view, we only use a certain segment of our belief systems at any given time: the one that allows us to act successfully in the circumstances at hand, or the one that answers to a given subject matter. By only having so many beliefs in any fragment, we can then explain inconsistencies in our system of beliefs, as well as its incompleteness. Understanding agents as being fragmented, or as having a fragmented belief system, then, comes from them being limited in one way or the other. If there were a God with unlimited computational capacities and unbounded by any of the earthly limitations that human agents face, then there is no reason to suppose that such a being has a fragmented belief system[2].

While being able to avoid (BC), the fragmentation strategy by itself does not avoid fragment-relative versions of either (LI), or (NE). To see why, first consider the proposal of how logical implication works in the propositions-as-sets-of-possible-worlds view. According to this approach, a proposition, *p*, logically implying that *q* just is for *p* to be a subset of *q*, or, equivalently, for *q* to be a superset of *p*. If an agent believes in one fragment that *p*, then for all they believe, the actual world is, for them, one of the possible worlds within *p*. Since *p* implies that *q*, then every world within *p* is also a world within *q*, and so the agent also believes in *q* according to the fragment in which they believe that *p*. This principle is also often taken to be unacceptable. Even when we have a small number of beliefs, it does not seem to be the case that we believe everything that logically follows from them in any given belief state. For instance, let us say that we only have three atomic beliefs in any given fragment - A, B and C -, plus everything that follows logically from them. If an agent believes that A, B and C, then the agent should also believe, given (LI), that $((A\&B) \rightarrow C)) \rightarrow ((A \rightarrow C) \vee (B \rightarrow C))$. However, this last

[2] See Yalcin (forthcoming) for a differing view. Here it is merely of interest to note that fragmentist theses often find their support in the fact that they are supposed to offer a plausible account of bounded or minimal rationality that applies to human agents, as opposed to the standards of ideal rationality for logically omniscient beings.

formula expresses a complicated relation given by the logical constants involved, and it is not a trivial matter to come to believe that it is true simply by believing A, B and C: students often seem to fail to believe it when they take a Logic course for the first time. In fact, Priest (2012, p. 14) for instance thinks that the formula is subject to counterexamples, such as the following: if (if you press switch A and switch B, the light will go on), then it's either the case that (if you press switch A, the light will go on) or that (if you press switch B, the light will go on). Further, in this view the agent should believe every necessary truth according to any of their fragments, for in the possible worlds view of what logical consequence comes down to, any set of propositions logically implies a necessary proposition[3]. As for the fragment-relative variation of (NE), it suffices to note that it is still the case that necessarily equivalent propositions are being treated as being one and the same proposition, so that in any one fragment in which an agent believes one proposition, they also believe the other.

The fragment-relative variation of (LI), it was just argued, should not be accepted. As for the variation for (NE), it seems that it is even more damaging than its fragment-independent version: even if (NE) fails and believing, say, that *5 + 7 = 12* is not the same as believing some other necessary proposition, like *Euler's Identity*, it seems even worse to accept that in every fragment in which one believes one of them, one also believes the other[4].

The fragmentist strategy, while being able to accommodate some inconsistencies in an agent's system of belief, is not capable of avoiding principles that still require too much of minimally rational agents. So, if we only add fragmentation to the possible worlds picture of content, we would have to come to an opposite conclusion to that of Lewis, (1986, p. 34): with *possibilia* and fragmentation we *are* still left with a problem of logical omniscience. While Stalnaker's (1984) and Lewis's (1986) picture can provide an explanation for some inconsistencies in our system of belief, as well as its incompleteness, they cannot provide a way to deal with *all* such phenomena.

Fragmentation, however, is not the only move available for one that proposes to explain mental and linguistic content in terms of possible worlds. Our criticism so far applies only to such a strategy, and in so far as it is the only one at play. In a recent paper, Yalcin (2018) proposes to take the project further, influenced by Yablo's (2014) work on subject matters. According to this elaboration of the current proposal to deal with the problem of logical omniscience, beliefs are sensitive to questions or subject matters, which are (for Yalcin) modelled as partitions of logical space. The gist of the idea runs as follows: when considering a

---

[3] See Beall and Restall (2006) for a more complete description, discussion and criticism of this way of conceiving of logical consequence.

[4] Here the basic point is simply that the fragment-independent variation of (NE) at least leaves open the way out that while an agent that believes in one, must believe in both propositions, the fragment in which they believe in one or the other, or both, can be different. The fragment-relative principle (which the view here expressed has to accept) has to claim both (as well as their conjunction) are believed in every fragment.

given question or topic, some distinctions between possible worlds become visible for the agents considering them - the ones that correspond to divisions in the partition - while some other distinctions stay hidden in the background; what the agent believes or knows is then taken to be one of those sets of worlds that correspond exactly to a union of cells of the partition.

The upshot of this elaboration is that, as stated, the fragment-relative variation of (LI) fails. In a given fragment, while *p* is believed, and logically implies *q*, *q* might not be believed by the agent, as it may not, unlike *p* does *per hypothesis*, correspond to a union of cells of the partition. Here it is of note that while the closure principle so far considered fails, there is still one that holds:

Closure under visible logical implication

(VLI) If (i) A believes that *p* in a given fragment, *F*, and relative to a question, *Q*, (ii) *p* logically implies that *q*, and (iii) *q* is visible relative to *Q*, in fragment *F*, then A believes that *q*.

Here, Yalcin (2018) maintains that (VLI), as opposed to the basic and fragment-relative varieties of (LI), is an acceptable principle. This closure principle is discussed later on. Taking now for granted that (VLI) is acceptable, what should we say about (NE)? If propositions are just sets of worlds, then if any proposition, *p*, is visible relative to a given question, so is *q*, for any proposition *q* that is necessarily equivalent to *p*. Adding question-sensitivity, then, doesn't seem to get us anywhere in regards to (NE) – a version of the principle in which it is further added a relativisation to questions alongside the already introduced relativisation to fragments would yield the exact same results as the fragment-relative version.

Yalcin (2018, pp. 35-37) also provides a sketch of a solution for this last problem. Yalcin's suggestion, as well as alternatives to it will be considered in an effort to preserve a possible worlds account of mental content that does not fall prey to the problem of logical omniscience. Fragmentation dealt with closure under believed conjunction; question-sensitivity (or so it is claimed) deals with all troublesome varieties of closure under logical implication. Is there any other feature we might add to deal with closure under necessary equivalence within a possible worlds framework?

In the first part of the present dissertation, a reply will be considered that readily suggests itself if the answer to the question of if there is anything one might add to fragmentation and question-sensitivity to block (NE) is negative: to go beyond possible worlds and introduce impossible worlds in our model for notions such as belief and knowledge. Then in the second part we consider various proposals based on the notion of a subject-matter, and finally provide a tentative solution for the problem of logical omniscience that makes use of an extended space of worlds, as well as a dynamic epistemic operator defined in terms of what questions an agent considers.

PART I

-

IMPOSSIBLE WORLDS

Given the supposed failure of possible worlds accounts of notions such as belief, knowledge, as well as postmodal notions such as grounding or essence (if Fine (1994) is right in his critique of analyses of essence in terms of properties an object has in all possible worlds in which it exists), a solution that naturally comes to mind is to extend our theoretical framework to include impossible as well as possible worlds. As we saw in the previous section, possible worlds views of all of these concepts seem to run, *mutatis mutandis*, into a common issue: with possible worlds alone, we can't make sufficiently fine distinctions, namely we can't distinguish necessarily equivalent propositions. According to the conception of propositions as sets of possible worlds (the ones in which they are true), necessarily equivalent propositions are one and the same, since they are true in the exact same possible worlds. However, there seem to be instances where we can distinguish between propositions that are true in the exact same possible worlds. Let us suppose that one afternoon while drinking water, Lavoisier believed himself to be doing just that, drinking water, which we can express by stating that Lavoisier believed that *Lavoisier is drinking water*. However, it is also an historical fact that when the chemical structure of water molecules was discovered, Lavoisier had already passed away. It seems, then, that we could not truly attribute to Lavoisier, in that afternoon, the belief that *Lavoisier is drinking H2O*. Even if he could, based on his knowledge of the chemical elements, come to have the concept H2O, it seems implausible that he held the belief that water had this structure. And here comes the difficulty: if propositions just were the sets of possible worlds in which they are true, then *Lavoisier is drinking water* and *Lavoisier is drinking H2O* would be the same proposition, since every possible world in which one is true is also a world where the other is. It seems true, however, as was just seen, to attribute belief in one of them, but not the other, to Lavoisier, in which case there must be some difference between them. If this is right, propositions, then, are not simply sets of possible worlds.

By making room in our ontology for impossible as well as possible worlds, an immediate reply becomes available. While true in the same possible worlds, *Lavoisier is drinking water* and *Lavoisier is drinking H2O* are not true in the same impossible worlds. Assuming, as we have so far, that water is of necessity identical to H2O, it would be impossible for Lavoisier to be drinking water but not H2O. And that is all, according to some proposals, we need to have an impossible world that represents Lavoisier as drinking water but that does not represent him as drinking H2O at the same time. As a general principle, proponents of this particular way out accept the following:

(NP) If *p* is impossible, then there is an impossible world that represents that *p*.

This principle, often referred to as *Nolan's Principle*, goes hand-in-hand with the common corresponding principle for possibility:

(P) If *p* is possible, then there is a possible world in which *p*.

While they have gained very wide-spread acceptance, possible worlds have had to prove their ground, as initially there were doubts regarding whether the notion itself was meaningful, about their queer nature, and about 'involvement with modal talk'[5]. Similarly, there are a number of reasons philosohers have proposed to not accept impossible worlds. If *possibilia* were already a hard pill to swallow in terms of extending our ontologies, accommodating for *impossibilia* seems to be too hard of a task to accomplish in a room that many would already deem too crowded. In the next two paragraphs, some questions that have been raised regarding the usefulness (or lack thereof) and viability of the notion of an impossible world are presented.

It has seemed to many, for instance, that *surely* even if impossible worlds are to be accepted in some way or another, there are not so many impossible worlds as to make room for every impossibility[6]. Our semantics should give us a compositional theory of meaning, and with so many impossible worlds thrown into the mix, we lose compositionality. Here's why: it is impossible that both A and B are the case, but not A&B; the meaning of "A&B", is given by compositionality from the meaning of the conjuncts, namely, it seems that the meaning of the logical constant "&" consists in that if the terms on each side are true, then the resulting formula is also true; but since there are, by (NP), enough impossible worlds to represent every impossibility, there is at least one impossible world in which A and B are true, but not their conjunction, we lose a way to determine, based on the meaning of the conjuncts, the meaning of the conjunction: we lose compositionality. It seems, then, that we should not allow for any impossibility to have a corresponding impossible world. If we give up (NP), however, how are we to be sure that there is indeed an impossible world in which Lavoisier drinks water, but does not drink H2O? Is there any principled way to distinguish between impossibilities, or should we find a way to salvage (NP) and have all impossibilities be counted in? And even if there is such a distinction, how to decide which group of impossibilities are to be represented by impossible worlds? With the move to impossible worlds, can we keep understanding propositions as sets of worlds? If not, do we not lose one of the major reasons to engage in talk of worlds in general, possible or impossible?

Closely related to these worries, there is a further question about what impossible worlds are, about the *nature* and existence of impossible worlds. Do impossible worlds exist? If so, do they work, as in Lewis (1986) possible worlds do, as restricted quantifiers over what there is, and are we then to conclude that we can get rid of the restriction on the quantifiers and, by accepting impossible worlds, get contradictions *simpliciter*[7]? Are impossible worlds entities of the same kind as possible worlds? If not, what are they?

All these questions are puzzling, and it may even seem mysterious how one should proceed in beginning to tackle them. Hopefully, the first part of the present dissertation provides some clarity on how some of these can and have been addressed in the literature, as well as to circumscribe a number of especially promising hypotheses. In what immediately follows, rival

[5] Quine (1953).

[6] See for instance, Stalnaker (1996, p. 201).

[7] This is the famous argument presented in Lewis (1986, p. 7, n. 3), against impossible worlds based on the Principle of Exportation. For this reason, we will accept below that impossible worlds should not be genuine.

conceptions of the nature of *possible* worlds are examined, which will lead us on to an interlude dedicated to analysing various forms of ersatzism. This has the purpose of enabling us to assess which conceptions of possible worlds are most promising for, with due changes, modelling impossible worlds, as well as to weigh the merits and objectionable elements of those that could more plausibly fulfil the role. From this discussion it will become apparent that some options are much more promising, both on their own accord, and as ways to model impossible worlds. In the latter sections of this part of the dissertation, concerns regarding the usefulness of those more promising ways of understanding impossible worlds will be considered. By the end of it, it will come to light what, if any, conceptions of impossible worlds are coherent and theoretically useful.

The nature of possible worlds: genuine or ersatz?

In order to understand what a world is, a crucial feature to get a grasp on is how worlds represent[8]. For instance, if we say that a given proposition, *p*, is possible if it is true in some possible worlds, then similarly we might say that some possible worlds represent reality as being such that *p*. But how do worlds represent?

According to the account of possible worlds favoured by David Lewis (1986), a given world represents what it does by containing what it represents as a part of itself. A possible world, *w*, represents that something is *F* by containing at least an object that is *F*. On this account, merely possible worlds have as much reality as our own, and so do the objects contained in them. According to Lewis, they are as concrete as our own, but just happen to be spatially and temporarily separate from us and, from our perspective, to be merely possible. In this view, for x to be spatiotemporally separated from all the objects that are part of a given world, *w*, entails that, from the point of view of *w*, x is a merely possible existent, meaning something that exists but isn't actual.

Lewis's account has many interesting, if not startling, features. Here, however, it suffices to analyse the more general claim that worlds represent that so-and-so by containing the relevant objects instantiating certain properties and relations. Following Berto and Jago (2019), let us call such worlds *genuine*. In contrast, we can define a family of *ersatz* conceptions of worlds, according to which worlds represent in some way other than by containing the required objects and properties as parts. In this very non-committal way, *ersatzism* is defined by denying that worlds are genuine. Ersatzism then, on this definition, encompasses quite different perspectives on the nature of possible worlds.

[8] Stalnaker (2012) has argued that worlds (which he understands as maximal properties of the total universe) do not *represent*, just like properties in general do not represent what they are applied to, but rather are simply instantiated or not by their corresponding objects. For now it is of use to take note of the fact that whatever relation possible worlds bear to what they "claim" is true, it is in need of further explanation. If it is a question how exactly and why worlds represent what they do, it is also a question how worlds come to be the properties that they are, and why the universe instantiates one of them and not any of the others.

A complete discussion of the merits and failures of each alternative when applied to *possible* worlds would lead us astray from our present goal. However, some points in the debate carry over to the discussion about the nature of *impossible* worlds. Therefore, some of the questions raised in the debate on the representational nature of possible worlds are here explored.

Reduction of modality and the incredulous stare

Lewis's understanding of possible worlds, while supposedly having a lot going for it, has been met with strong resistance by a substantial part of the philosophical community. Possible worlds, as mentioned before, seem to be *prima facie* entities that are hard to accommodate in one's ontology. The case for possible worlds, however, seems to get even more unintuitive when it is further claimed that they are just as concrete and exist in the same way as the very own world we inhabit. Why should one believe such a tale? One of the main reasons for it seems to just be the usefulness of genuine possible worlds in theoretical practice[9]. Lewis-style possible worlds would then be entities that, just like electrons, for instance, we must accept as existing, given the requirements of our best comprehensive theory of reality.

One way in which genuine possible worlds seem to be theoretically advantageous in contrast with other conceptions of possible worlds is that they provide a complete reduction of modal talk to talk of possible worlds. Let us go back to (P). In it, the notion of something being true in a possible world (expressed by «there is a possible world, *w*, in which *p*») gives us an explanation of why it is true. But how then are we to understand what it means to be true in a possible world? According to David Lewis, a given proposition being true in a world just is for that world to contain as a part the relevant objects instantiating certain properties and relations. What from the standpoint of the actual world are merely possible objects have just as much reality as the objects of our own world. We have then provided, with this conception of what a possible world is, a complete reduction of modal talk to quantification over possible worlds,

[9] Lewis (1986, pp. 3-5); Jacinto (2013, p. 7). The other argument (see Lewis (1973, p. 85) in favour of realism about possible worlds - the argument from familiarity - is that in ordinary talk we say that there are «ways that the world can be». This talk is then taken at face value as quantification over such things as ways the world can be, or, in other words, possible worlds. As Stalnaker (1976, p. 66) points out, the more precise argument seems to be that there is a presumption to take ordinary talk at face value, and for us to not take it so, we must show that taking it that way would lead to difficulties of some kind. The burden is then on the possible worlds sceptic to say what problems does taking it at face value lead to, which Lewis thinks cannot be done. But while we might take this ordinary talk at face value, we need not accept automatically that these ways the world can be are as Lewis takes them to be: it invites us to accept a *realist* conclusion, not necessarily an *extreme realist* conclusion. As Stalnaker (1976) shows, we can be moderate realists simply by accepting «ways the world can be», while rejecting two of the stronger theses advanced by Lewis in regards to the nature of possible worlds: that there is no fact of the matter as to which world is absolutely actual, as opposed to actual from its own point of view; and that worlds be taken as *concrete* irreducible objects in their own right, with equally concrete parts. Further, it is perhaps the case, as Stalnaker (1976, p. 68) claims, that this second argument of Lewis in favour of the existence of concrete possible worlds ends up conflating certain properties (the «ways the world can be») with the objects that would instantiate them («I and all my surroundings»). What this argument shows is that we should accept possible worlds, in so far as they are properties that would be instantiated by the total universe, not that there are actually concrete objects of a certain kind that instantiate them.

which are concrete objects, and quantification over the objects that are mereological parts of the possible worlds, as well as instantiation of properties and relations by those same objects, neither of which are mysterious.

The gain of providing a complete reduction of modal talk, however, can only count in favour of Lewis's extreme realism against rival conceptions of possible worlds if no other conception with less extreme ontological commitments can also provide such a reduction. Lewis goes on to show that three varieties of ersatzism that he identifies as interesting proposals to argue against end up failing as good alternatives to his own realistic view of possible worlds, and that all of them have to accept primitive modal talk. Given the scope of the present reflection, I will ignore the discussion on Pictorial Ersatzism, I take Lewis's (1986, pp.165-174) arguments to be decisive against this proposal. Henceforth, the remaining varieties of ersatzism are discussed[10].

Furthermore, it is not an undisputed claim that reduction of modal talk to quantification over objects and instantiation of properties and relations is even desirable. If Stalnaker (2012) is right, a theory that, like Lewis's, attempts to present such a complete reduction of modal talk should be regarded with suspicion, in the same way one would regard wholly reductive accounts of such notions as existence and quantification as suspicious. More would have to be said on whether such a reduction should be pursued, however given the number of questions that had to be considered, this is left as an open question to address in future work.

[10] Another way in which the forthcoming discussion is incomplete is that fictionalism in respect to possible worlds is not considered. There is, however, a reason for this absence, one that makes this omission, I hope, more acceptable. The end goal of the present considerations is to consider whether there is any way of making sense of the notion of an impossible world that is theoretically acceptable (never mind for now questions of usefulness). Genuine impossible worlds are a hard bargain, and so we have to do with some other understanding of what impossible worlds are. For the purposes that impossible worlds are supposed to do, it will be seen that these are best understood as constructions of some kind (out of propositions, or states of affairs, or sentences of a worldmaking language). If this is right, then whether impossible world talk is fictitious or not will come down to whether the elements they are constructed out of are themselves to be taken as fictitious. If, say, proposition talk is to be taken as fictitious, then impossible world talk, if impossible worlds are taken as constructions out of propositions, is to also be taken as fictitious. Whether impossible worlds are themselves fictitious objects or not will be considered in future work. For now, what matters, for future discussion, is to show that extreme realism is not the only satisfactory view of the nature of possible worlds.

Magical Ersatzism

Ersatz and genuine possible worlds are distinguished by the way they represent: genuine possible worlds represent that *a* is *F* by containing as part an object *a* with property *F*; ersatz worlds don't. In what way do ersatz worlds represent, then? That will turn on the specific flavour of ersatzism one adopts. The varieties are then individuated by what claims they make about how worlds represent.

We start by looking at a view which refuses to say more about how ersatz worlds represent what they do, that they do so is taken as an unanalysable primitive. Lewis (1986, p. 141) refers to this position (pejoratively) as *Magical Ersatzism*, claiming later in Lewis (1986, p. 178) that it seems to be only in a way not short of magic that one can understand the proposed notion of primitive representation. In the discussion that follows we'll be using "Primitivism" (following Berto and Jago (2019)) and "Magical Ersatzism" (following Lewis (1986)), interchangeably to designate this form of ersatzism.

Before going on to consider Lewis's criticism of this family of positions, let us say more about what Magical Ersatzism tells us about worldly representation, what its overall theoretical framework is like, and what the primary motivation for it seems to be. In this picture, worlds are a subset of simple abstract elements in the best theory of modality, that is, they lack any structure, and they are taken as primitives. If worlds are just a subset of these elements, what defines them? That they are *maximal*: for every other element, a world either implies it or its negation (this is an imprecise way of speaking, but that we can make clearer in the following way: given an element that represents that *P*, the negation would be the element representing that ~*P*). An example might make clear how this is all supposed to work. Let us consider the possibility that a donkey talks. In this view, it being possible that a donkey talks is for there to be an element that represents a donkey as talking. This element, let us assume, would represent nothing else besides that a donkey talks, and it would then be implied by all the maximal elements (*i.e.* all the possible worlds) in which a donkey talks; and that is why all and only these worlds are ways the world can be in which a donkey talks. To be a possible world that represents that *P* is to be one of those maximal elements that implies the element that represents that *P*.

Given the exposure so far, one would justifiably be unsatisfied with how specific this view is. There is the actual world (in Lewis's sense of a maximal concrete object that is the mereological sum of all objects spatiotemporally related in it) or, in some ersatzists' preferred terminology, the *total universe*[11], and there is what we might call, following Lewis (1986, p.

[11] This distinction is important, as ersatzists don't take in general worlds to be concrete, and that includes the actual world. For instance, in Stalnaker's (2012) view, worlds are properties that the total universe (*the only* Lewis-style world that exists, see van Inwagen (1986)) can have, and that includes the property that the total universe instantiates. The actual world is precisely this latter property, being the fact that it is instantiated what distinguishes it from the other maximal properties. The total universe is then a concrete world in the way Lewis understands them, and the actual world is the maximal property the total universe instantiates.

174), a selection relation – a relation that maps certain happenings in the total universe to the elements that represent them. However, what *are* these elements? And what is this selection relation like?

Before getting into these questions, let us consider briefly what seems to be one of the main motivations for Magical Ersatzism that comes directly from the way semantics is done. In Kripke models, standard in possible world semantics, worlds are points which are then stipulated to represent certain propositions as being true and others as being false and to be in certain accessibility relations to one another (and, in systems of modal logic as strong as, or stronger than, T, also to themselves). Since this can be done without any further specification of what worlds are and how they represent, and if, as the magical ersatzist claims, various ways of conceiving of how worlds represent end up having problematic features, why not give up on giving an answer to the question of how worlds represent? Further, perhaps there isn't even a meaningful, or at any rate useful, question to be asked here, perhaps only representation is needed, and how exactly it works is beside the point.

When doing formal semantics this primitive sort of ersatzism seems to work, and we don't seem to need to make any further claims about what worlds are and how they represent, save for those inherent to our formal system. However, how should one proceed when discussing the metaphysics of possible worlds? As a first approximation, it seems to at least be coherent to ask how in general a possible world represents that there is a talking donkey. Given this, in the context of Magical Ersatzism it also seems to be appropriate to ask: why does *this* element represent that a donkey talks and not, say, that a cat is on the mat? Worlds in the primitivist view are simple, unstructured, abstract entities that play the role required of them by given models. But for this to be the case, it seems that more needs to be said about how to pick certain elements to play each of the roles required of them. For instance, why is it a given set of worlds that represents that a cat is on the mat, and not any other set? The first answer, as noted, is simply that they all are maximal elements which imply the element that represents that a cat is on the mat. But what allows us to *select* that given element to do that job and not any other? For instance, why are such-and-such elements the maximal ones and not some others?

Lewis's dilemma for Magical Ersatzism

Lewis's argument against Magical Ersatzism comes down to a dilemma of what the nature of this selection relation is[12]. The selection relation[13] ought to be either an internal relation, i.e., a relation that holds in virtue of the intrinsic properties of the *relata* taken individually; or an external relation, i.e., a relation that holds in virtue of the properties of the

[12] There is a broader argument against Magical Ersatzism, which comes down to Lewis's own methodology, and which he applies to other perspectives, such as the view that there are structural universals. See Fisher (2015) for a presentation and discussion of this broader argument. In the end of the present section we will come back to it.

[13] Here I am ignoring extrinsic relations and focusing on internal and external *intrinsic* relations. This is so because it is very implausible that the selection relation would be extrinsic. See Lewis (1986, p.182), Jubien (1991), Zaragoza (2007) and Fisher (2015) for supporting arguments.

sum of the *relata*, but not in virtue of the properties each of them has separately[14]. As an example, *having-the-same-mass-as* is often taken to be an internal relation, given that it holds in virtue of the intrinsic properties of the related objects (in specific, the property each of them has of having a given amount of mass), so that any duplicates of the *relata*, *i.e.* with the same (relevant) intrinsic properties, will be in the same relation to each other. As for external relations, the paradigmatic examples are spatiotemporal relations, such as *being-one-meter-away-from*. That I am one meter away from my bottle of water in no way depends on my nature or my water bottle's nature, but only on the properties of me and my water bottle taken together (for instance I might have the property of being at the spatial location *s1*, my water bottle at the spatial location *s2* – which are extrinsic properties of me and the water bottle, respectively – and taken together these properties imply that we're at a given distance, given the distances between *s1* and *s2*).

In what of these categories should the selection function fall on? Let us consider first the hypothesis that it is an external relation. Then it seems that whatever simple abstract elements the happenings in the total universe pick up, they could have picked some others. However, this seems wrong, if a donkey talks, then it seems that the maximal element picked by the talking donkey in the total universe must be one according to which a donkey talks, and not any such element will do the job. An element that represents that a donkey talks is not the same as one according to which there are no donkeys, and yet either of these could, if the relation was external, fulfil the required role.

Here the primitivist could reply that indeed any element could play the desired role: that would just be a situation where the element that actually represents that there are no donkeys would represent something else, namely that a donkey talks, and some other element would be selected to do the work of representing that there are no donkeys. It is indifferent what element are selected, and even what, to some extent, their properties are, provided there are enough of them. This echoes a more general structuralist approach. Just as some philosophers of mathematics claim that any series of objects will do to be series of the natural numbers, provided it is structured in a certain way and there are enough objects in it, so perhaps some possible world theorists may claim that we can take any set of objects to be the possible worlds, provided there are enough of them and that they are maximal (which is to say, that we also pick some other objects and that we stipulate that a certain relational space ties them in the appropriate way to make all the implication relations we want there to be to hold)[15].

A potential problem with this way out is that it invites a certain eliminativist attitude. Just like some structuralist views of what natural numbers are invite us to think that there is nothing

[14] This peculiar characterization is present in Lewis (1986, p. 176). Some pages later, in Lewis (1986, p. 180), a different, more common way of characterizing external relations is assumed without argument to be equivalent to this one: an external relation is one that holds independently of the natures of the *relata* (that is, their intrinsic properties). See Jubien (1991) for a critique of this assumption.

[15] We will come back to a similar argument proposed by Denby (2006) by the end of the section.

about them (or at least that is of interest for our theoretical purposes) besides their structural features, so we can say the same about possible worlds: it doesn't matter what they *are*, as long as they are structured in a certain way[16]. And again, while this might be the case when we're simply doing formal semantics, it seems there is a need for a more satisfying answer when we engage in metaphysical debates and want to say more about the nature of possible worlds. In fact, if this reply was right, why would possible worlds even need to be abstract? Perhaps the reason is that there aren't enough concrete objects to do the job, but even if so, why not take possible worlds to be a hybrid set of concrete and other, familiar, abstract objects? If it doesn't matter at all what they are, as long as the relational features are adequate for the job at hand, why introduce a new set of simple abstract objects besides the concrete and abstract objects already available? This seems to leave us with only "so-called" possible worlds, and not really with a new set of objects. We have then eliminated possible worlds from our theory, and we have merely kept the *ways-things-can-be* role to be fulfilled by whatever we may so choose. This reply to the problem, then, appears to only be available for uses of possible worlds semantics when constructing formal models, for it seems to be only then that it is completely indifferent (within the aforementioned restraints) what the possible worlds are taken to be.

Given the failure, for current purposes, of the reply just considered, Lewis seems to be right to insist that this purported external relation is untenable. It seems to be in virtue of the total universe's intrinsic nature (namely, that it contains certain objects as parts, instantiating certain properties) that some elements are selected. It is because the total universe has a talking donkey that it picks up the element that represents it, whatever that element might be. In fact, it seems to be because of this that we are allowed to say that the elements are representational in the first place: elements are representational so long as they map onto some intrinsic properties of the total universe. By the definition of an external relation, however, the selection relation should hold independently of the intrinsic properties of the *relata*. And this contradicts the fact that whatever elements are picked, depends in part on what the intrinsic properties of the total universe are. The selection relation, therefore, cannot be an external relation.

The only option left is that the selection relation is an internal relation that holds in virtue of the intrinsic properties of each of the *relata*. But what *are* the intrinsic properties of the elements? Given that these elements are abstract, how can we have access to them[17]?

[16] Lewis (1986, p.179)

[17] It has also been pointed as a criticism to Lewis's (1986) perspective that his possible worlds, being causally disconnected from us, are likewise inaccessible, so that in his theory we wouldn't be able to attain knowledge of facts about modality. As stated, however, the genuine possible worlds of Lewis's theory are not, given the principle of exportation, adequate to be expanded to model impossible worlds. Here we are merely interested in discussing ways of understanding possible worlds that, with certain changes, could give us a plausible account of the nature of impossible worlds. The argumentation, then, follows without us having to get into this debate. It shall be noted, nonetheless, that Lewis's criticism of the option of taking the selection relation to be internal isn't simply that the abstract elements are epistemically inaccessible from us, but rather that for us to grasp what the maximal elements are, we would need to be able to grasp an infinite pairing of abstract elements' natures with happenings in the total universe, given that only such a pairing would allow us to individuate a complete and maximally specific possible world on this conception. We are, however, claims Lewis, unable to grasp such a pairing.

Furthermore, given the number of elements needed to represent every possibility, we need to have an infinite number of said intrinsic properties of the abstract elements, otherwise we wouldn't be able to make sufficient distinctions between the different possible selection relations. If the total universe were in a way such that it contained a talking donkey, then (assuming, as we are now, that the selection relation is internal) there must be an element with a certain intrinsic property so as to be that element the one that represents that there is a talking donkey. Similarly, intrinsic properties of this sort would have to be presented for every distinct possibility. Given that possible worlds are maximal, we would have to grasp an infinite conjunction of associated properties of objects in the total universe and abstract elements for us to grasp what the relation would be between a given possible world and a total universe. Namely, it seems we would need an infinite pairing of intrinsic properties of worldly objects and abstract elements to get to grasp the maximal element that entails them. This, however, we can't grasp. The selection relation cannot, therefore, be an internal relation.

Since we can't understand the selection relation as an external relation, given the work it is supposed to do, and likewise we can't grasp it if it's an internal relation, there is no way Magical Ersatzism can be successfully understood, and if its proposers claim they are able to, they must do so in a way not short of magic.

Dealing with the dilemma: the relation is external

A first way that the Magical Ersatzist might want to avoid the dilemma is to insist on a different way to uphold that the selection relation between the total universe and the abstract simples is external.

Zaragoza (2007) starts by noting that Lewis's argument relies on a recombination principle that is very strong (that is, allows for a lot of recombinations, and so reduces the number of necessary connections between objects): the principle that two independent coexisting objects *x* and *y* might be in any external relation with each other, as well as to every other independently existing object. This principle, it seems to me, is very intuitive, and seems to come out true directly from how external relations are commonly defined: they are those relations that hold independently of the nature of their *relata*. Given this definition, a certain external relation holding is wholly independent of any object's nature, and so it would seem strange that, even so, certain necessary connections held between the objects related by it. If this necessary connection is not grounded in the nature of the *relata*, in what would it be?

It is not this principle, however, that Zaragoza calls into question, but rather Lewis's ability to uphold it. In particular, Zaragoza mentions two examples that Lewis seems to have trouble accommodating: the possibility that there would be nothing; and "island universes", that is, spatiotemporally isolated, but interrelated universes[18] that belong to the same possible world (understood in Lewis's sense as a maximal concrete object).

In the present context, this line of reply is not very promising, however. For starters, it is of note that it consists in a *tu quoque*. If Lewis's perspective suffers, as well as the magical erstatzist's, from problems in upholding the aforementioned recombination principle, then so much the worse for Lewis and the primitivist, we should look for a different perspective that allows us to keep holding to the principle. Secondly, perhaps one can retain much from Lewis's perspective and still be able to accommodate for the possibility of "island universes", perhaps appealing to some other external relation holding between the disconnected spacetimes (such as a primitive worldmate relation). And thirdly, it is controversial that two worldmates can belong to interrelated but disconnected spacetimes, or at any rate that the notion is useful, with Lewis himself having presented "substitutes" (Bricker (2001), Lewis (1986, p. 72)) that his theory can countenance to do the work that island universes seem to do in a wide range of cases. For instance, Lewis can provide a basis for the hypothesis from brane cosmology for why gravity is so much weaker than the other fundamental physical forces: "The spacetime of the big world might have an extra dimension. The world-like parts might then be spread out along this extra dimension, like a stack of flatlands in three-space." (Lewis, 1986, p. 72). Gravity would then just be a force operating across spacetime along this extra dimension, on which the branes are "stacked", while the other fundamental interactions would all be intra-brane interactions. If the substitutes are enough to cover all cases we want them to cover is, however, up for debate. For instance, there seems to be a hypothesis left open by General Relativity that entails the existence of spatiotemporally isolated, but actual in respect to the same possible world, spacetimes. If that is right, then it presents a difficulty for Lewis's own proposal. But again, the objective here is not to defend Lewis from criticism like this, just to assess the merits of the Primitivist position.

Finally, as for the possibility that there is nothing, we might have different positions in mind. I'll highlight two: the possibility of there being nothing at all, and the possibility of there being nothing concrete. In regards to the former, it is at least disputable that there is indeed such a possibility. For instance, it seems to many that if abstract objects exist in any given possible world, that they must exist in all of them - after all, what would be the reason for them not to exist in the other worlds? Even for second-order contingentists, like Stalnaker (2012), the necessary proposition (one of the zero-placed properties of the total universe) exists in every possible world, since what distinguishes the possible worlds is what partitions of logical space (which correspond to different possibilities from the perspective of that world) exist for them, whereas logical space (which corresponds to the partition introduced by the necessary proposition) exists from the perspective of every world. It might be even the case that it is not possible for there to not be any individuals whatsoever. Let us accept with Stalnaker (2012) that singular propositions exist only if certain objects exist (for instance, the proposition expressed by

[18] This corresponds to Bricker's (2001) weak notion of an island universe. Bricker goes on to maintain that there are also island universes understood in a stronger sense as being totally isolated from one another.

“Socrates is a philosopher” does not have Socrates himself as a constituent, but nonetheless only exists in possible worlds where Socrates exists). Let us now consider the hypothesis that at least some necessary mathematical propositions are singular, i.e., their existence depends on the existence of individuals in a given domain. For instance, Lewis (1991) accepts that the empty set is the mereological sum of all individuals, which is itself an individual, so that all the pure classes are constructed from an individual. If his view is correct, then some propositions of mathematics will be singular. Being necessary, these propositions are, on this coarse understanding of what a proposition is, one and the same: the proposition that corresponds to logical space. Given that logical space exists from the point of view of every possible world, these propositions exist in every possible world. These considerations do not resolve the issue against the possibility of there being nothing at all or no individuals: they only have the purpose of showing that it is controversial that there is such a possibility, and that even second order contingentists like Stalnaker will not be able to uphold it.

I believe, however, that there is a further argument against the existence of the possibility of there not existing anything at all which does not depend on any of the views just considered. Regardless of what entities we take possible worlds to be, possible worlds play the role of maximal elements, as mentioned beforehand, that is, they correspond to maximally specific ways that the world could have been, representing for every proposition whether it is or isn’t the case. But a situation representing nothing at all, which would correspond to the putative possible world where nothing exists would not be a maximally specific situation, for since it does not represent anything, it does not decide whether any propositions are the case or not. It might be said in response that the envisaged world is not a possible world that does not represent anything, but rather where nothing exists. But if nothing at all exists in a world, then it is not clear how the world might represent that there is nothing. For instance, if we take possible worlds to be certain maximal properties, then the possible world representing that there is nothing would correspond to the property the world would instantiate if there had not been anything at all, but if there existed nothing at all, then likewise the property would not have existed, so that nothing would instantiate any properties. Similar considerations apply for other perspectives on the nature of possible worlds. Here note that the claim is not that possible worlds must represent in the way genuine worlds do, by containing the relevant individuals and properties as parts. Rather, the claim is that since what a possible world represents is a complete alternative to actuality, then it ought to be an accurate representation of actuality had it been the actual world. The argument is then that the putative possibility of there being nothing violates this constraint, as if actuality had been like that, then there wouldn’t be anything, and therefore no representations, and finally no accurate representations. An alternative way of thinking of the constraint we’re imposing is that the actual world should always exist from its own standpoint, so that if other worlds had been the actual world, then they should exist from the point of view of that world, whereas if nothing had existed, no actual world would exist either, so that it seems contradictory to claim that this putative possible would then be actualized.

Even if we accept this constraint, it might be said that the fact that the possibility of there being nothing at all can't be captured by possible world semantics doesn't by itself mean that there isn't such a possibility, it could just prove to be a limitation of the conceptual schemes we are employing to deal with questions of modality. I believe, however, that this line of reply isn't very promising, and in particular because the argument can be run even from a very intuitive and non-committal way of understanding possible worlds (it is indifferent, for instance, in regards to whether they are concrete objects, properties, states of affairs, sets of sentences in a worldmaking language, etc.).

It might be thought to be outrageous to even consider that there is no possibility that there would be nothing at all. That feeling, however, can be easily explained away in a manner that will lead to agreement with the present point. One of the classic problems in Philosophy consists in how to answer the question of why there is something rather than nothing. But what kind of nothing? For instance, when in the Philosophy of Religion philosophers ask themselves why God decided to create something instead of not creating anything at all, that is not the same hypothesis as the one just considered. In fact, by formulating that question, the existence of God and of a situation (in which God is deciding whether or not to create anything else) is already assumed. To see how unusual the possibility of there being nothing at all would be, consider that in it there wouldn't be the possibility of anything existing. So if the possibility of there being nothing at all were to be actualized, it would have been impossible to exist anything at all, since there would also not be any possibilities, including itself (which as was just claimed, should not be accepted).

It is not, of course, the possibility of there not being anything at all that Lewis (1986, pp.73-4) struggles with, but rather with the possibility of there being nothing *concrete*. This seems to follow immediately from the modal realist's claim that possible worlds are just maximal concrete objects that all exist on the same level. I believe that this is a fair point, and a disadvantage of Lewis's perspective (that is, I accept that there is a possibility that nothing concrete exists). Lewis (1986, p. 73) accepts the conclusion that there is no world in which there is nothing concrete, and is happy to do so. I do not wish to take sides on this particular debate. Rather, I will just proceed to provide a way that one could, if they wished to claim that there is the possibility of nothing concrete existing whilst retaining most of Lewis's theoretical framework. I believe this possibility can be accommodated by introducing a singular non-concrete world in the extreme realist's model, coupled with a slight, but largely innocuous, change in the definition of concreteness. One would start by introducing *the* world that trivially satisfies the condition of being a spatiotemporal (or analogously-spatiotemporal) whole[19]. One could then write a condition for being a Lewisian possible world in the following way: if *w* is a possible world, then for every x and y in *w*, x is spatiotemporally related to y, and there is no z

[19] This world would function in Lewis's mereology in a similar way to how the empty set works in pure set theory. While Lewis doesn't accept such a "null individual" (Lewis (1991, p. viii)), Harry C. Bunt (1985, p.54), who engages in a similar project to that of Lewis's in *Parts of Classes*, accepts an "empty ensemble", which, as the author states: "can be regarded as the empty set, and vice versa".

not in *w* such that any object in *w* is spatiotemporally related to *z*. A world where there is nothing would be a world that vacuously makes it true that every part of it is spatiotemporally related to every part of it and nothing else (this second condition is met as everything there is will be by definition outside of the domain of the empty world, and, so, disconnected to every object in it). Here the worry is that this world would itself be something concrete. As Lewis (1986, p. 73) puts it: “The world is not like a bottle that might hold no beer. The world *is* the totality of things it contains, so even if there's no beer, there’s still the bottle”. However that there is a bottle, *i.e.* a container, as was seen (and that’s why we went to such great lengths in discussing the possibility of there being nothing at all), is not itself problematic, the troublesome part is that this container is also taken to be concrete. But should it? If we were to think of the container as an empty spacetime, say, then it would seem so, but given that what we have is simply a world with where nothing concrete exists it is not as clear that we should still say that it is concrete. Perhaps we could use the following as a definition of concreteness: a given object is concrete if and only if it or one of its duplicates bears spatiotemporal or analogously spatiotemporal relations with other objects. In this definition, the world that trivially satisfies the condition of being a spatiotemporal whole would not be concrete, since neither it nor any of its duplicates bears any of the required external relations to other objects. We could then account, without changing too much of the original perspective, for the possibility of not existing *in a possible world*, any concrete objects[20]. If on the other hand we keep upholding that the empty world is itself concrete, perhaps it is not as problematic that it comes out true that there is no possibility of there being nothing concrete. We would be including among the cases where there is something concrete one in which we have an empty world. But this already seems like a deviation from most cases we usually consider when thinking about the possibility of there being nothing concrete, and so perhaps our intuition that there is such a possibility is based on other more familiar cases (such as there not being any matter or spacetime).

Zaragoza’s way out, then, doesn’t seem to be convincing. Let us go over the problems with it: firstly it doesn’t present an argument against the recombination principle Lewis appeals to, it just points out that Lewis, like the magical ersatzist, can’t uphold it, which does not weaken the objection against magical ersatzism per se; secondly, it is not clear that island universes are as threatening to Lewis’s position; and lastly, there doesn’t seem to be a coherent possibility of there not being anything at all, and, by tweaking Lewis’s modal realist position, we can account

[20] The discussion in the last paragraphs has been imprecise. Lewis, of course, can’t allow for the possibility of there not being anything concrete, if in stating that possibility we use the quantifiers unrestrictedly, for logical space exists from the point of view of every world, and since logical space is constituted by concrete worlds, then from the point of view of every world there exists (unrestrictedly) something concrete. I take this to be a disadvantage of his view, but I’ll ignore it for now. The possibility that this tweaked analysis allows for is simply for there to be a world where there actually is nothing. It corresponds, then, to what actually exists from the standpoint of a given world, not to what unrestrictedly exists. This possibility in turn seems to be enough to cover a range of cases that we seem to be thinking of when considering the possibility of there having been nothing, as when in thought we start from our own world and consider the possibility of there not having been a spacetime or fields, physical laws or any objects in external relations of this sort, and slowly start, in counterfactual reasoning, to take out everything existing in the actual world.

for one conception of there not being anything that seems able to account for several different scenarios we think about when considering the possibility of there not being anything. Even if Zaragoza's argumentation has proved to be lacking, can we go some other way in maintaining that the selection relation is external? Denby (2006) and Jubien (1991) give an affirmative answer, finding Lewis's argument unconvincing at every turn.

Denby (2006) starts off by providing a parody of Lewis's argument, giving as an example the *has-a-length-in-metres-of* relation between extended objects and numbers, which he takes to be external[21]. If we accept that it is an external relation, then it seems that we run into the same supposed trouble that Lewis identifies for the selection relation: namely it seems that what number gets picked up depends on the intrinsic property of the extended object of having a specific length. Denby (2006) also considers the hypothesis that this might be an internal relation, owing to also intrinsic properties of numbers. If every duplicate of a number (that is, every object sharing all the intrinsic properties with a number, but not its extrinsic properties) is that same number (that is, to be a given number is just to have those properties), then this condition is satisfied. But, even if the relation is internal, what the nature of this relation is seems to be an open question. In so far as it is, it seems that it shouldn't simply be settled by a general argument like Lewis's. This is the first moment of Denby's critique.

I believe that this criticism also fails. First, it is not clear that it should count against Lewis's argument that it has the consequence that such relations as *has-a-length-in-metres-of* are not external. If the recombination principle alluded to beforehand should be accepted, upholding it could be taken as a *desideratum* that any theory of external relations should meet. Since considering *has-a-length-in-metres-of* to be an external relation leads us to violate the principle, it seems to be an open question whether we should hold fast to the principle or give it up. That, however, will turn on how convincing the purported counterexample is. The point is that if we consider Lewis's argument under the light of the recombination principle, it ceases to be surprising why such a general argument could have the implication of settling the question of what some relations are. Secondly, perhaps the counterexample isn't as pressing as at first it might look like, and so it might not provide a good reason to reject the recombination principle. I agree with Denby that there is some pull to consider this as something other than an internal relation: it seems that there is a degree of arbitrariness in picking one number or another to represent the length of an object. However, should this be explained in terms of an external relation that somehow still depends on the intrinsic properties of the extended object? That, as seen, would lead us to accept necessary connections between independent objects that hold regardless of their intrinsic properties, which seems mysterious. If this arbitrariness can be explained in any other way that doesn't make us accept such a (I take it) unintuitive conclusion, we would be better off. A very ready solution seems to be open to us: that *has-a-length-in-metres-of* is an extrinsic, not an intrinsic relation, with the conventional or arbitrary part of the relation between the length and the natural number coming from the chosen standard for

[21] Denby (2006), p. 172.

transforming lengths into natural numbers – in this case, the standard defining a metre. An extrinsic relation is one that holds between the *relata* in virtue of their relations to other objects. Let us suppose, to make this example more manageable, that we're measuring a plank of wood. When we attribute to it, say, the number two, we are not simply in the presence of a relation between the plank and the natural number that would hold in virtue of the properties of the natural number and the plank taken together (recall that this is Lewis's definition of an external relation). Instead, it seems that we have in the background, so to speak, a relation holding between the plank of wood and paths of light travelling in the vacuum. This relation seems internal and would describe the ratio between the standard for what a metre is (how far light travels in the vacuum in 1/299792458 seconds) and the plank of wood. The length travelled by light in the vacuum in the specified amount of time would then be in an external and conventional relation to the number one[22]. With this we seem to have fully explained the pull to both considering the first relation as internal, and as external: it is internal in so far as the length of the plank of wood relates to the standard for the metre, and it is external in so far as it is conventional that the length of the path of light in the vacuum in the considered amount of time is associated with the number one. We should conclude, then, that the relation is extrinsic or internal, and still forego the hypothesis that it is external. This first argument against Lewis's objection is, then, defused.

The second moment of Denby's (2006) argumentation against Lewis seems to again try to show that, if the argument were right, it would be too general: it would affect all Leibnizian conceptions of modality that take the selection relation to be external. Denby considers a Leibnizian account of modality to be one that adopts the following schema: "it is possible that so and so iff there is a possible world that represents that so and so" (Denby (2006, p.162). All accounts of modality here analyzed are, in this sense, Leibnizian. As Denby (2006, p.172) notes, the argument against taking the selection relation to be external does not make reference to the simplicity or abstractness of the elements. It implies, then, that no Leibnizian account of modality can be appropriately grasped if they accept an external selection relation. As with the first moment of the argumentation, we will consider two lines of reply, namely if it indeed should count against the argument that it leads to this conclusion; and if the argument in fact applies to all Leibnizian accounts that rely on an external selection relation.

Denby (2006, p. 172) here echoes an argument of van Inwagen's (1986, pp. 207-210) to show that Lewis's own modal realism would be at odds with the argument presented against the magical ersatzist. Namely, Lewis accepts a relation of membership: propositions are sets of possible worlds and a given proposition being true in a possible world just is for that world to be

---

[22] Here it might be pointed out that the prime example of what an external relation is would then be considered to be an extrinsic relation, since it also seems to be presupposed in the *being-one-metre-away-from* relation that there is an intermediary relation to the standard for the metre at hand. Perhaps that is correct, but I believe it only shows a limitation in the language available for talking about lengths. In this case what we are referring to is the length that we conventionally refer to as "one metre", and we then say that the *relata* stand apart from each other in *that* length. And in being *that length* away from each other consists the external relation.

a member of that set of possible worlds. As van Inwagen (1986, p. 207-8) goes on to argue, however, the relation that holds between a set and its members is subject to the same argument that Lewis puts together against the magical ersatzist's selection relation. For one, the relation of membership can't be simply internal, for otherwise perfect duplicates would have to be members of all the same sets (see van Inwagen (1986, p.208)). But, let us suppose, electrons are duplicates of one another (the example is mine, adapted from Sider (2020, p. 66), not van Inwagen's), and yet belong to different sets: a given electron in my pocket is a member of the set of electrons in my pocket (supposing there is such a set); but the electrons in my glass of water aren't members of that set. What makes the electrons in my glass of water not belong to the set of electrons in my pocket isn't that they have any different intrinsic properties, one may assume, but rather the external relations they bear to other objects, namely spatiotemporal relations to my pocket. We have, then, to conclude that the membership relation is external. But here we are faced with the objection we are now considering: if the relation is external, how do we account for the supposed necessary connection between a member of a set and the set it is a member of? If the membership relation were external, then, by the recombination principle endorsed by Lewis, it should have been possible for the set to exist but not the individual, or the other way around, and any individual could belong to any set[23]. However, it seems that for instance the set {David Lewis} would not exist if David Lewis hadn't existed, meaning that also here we seem to find a necessary connection when confronted with an external relation.

Van Inwagen (1986) seems to think that since Lewis's argument would imply that we can't grasp the relation of set membership, and that in turn would mean that we don't understand ourselves when we do set-theory - which finally leads to confusion in classic mathematics - the argument must have gone wrong somewhere. Further, since the argument against Magical Ersatzism would be just as damaging against Lewis's theory, and since, van Inwagen claims, there is a *prima facie* case to accept ersatzism over extreme modal realism, that one should go the ersatzist route. I believe, however, that the conclusion is hasty. It is undoubtedly correct that set theory has proven to be very useful, but no matter how useful a theory turns out to be, an argument that would seem to raise issues for it shouldn't just be discredited for having the conclusion that one of the theory's core notions is not clearly understood, shall we not be able (as van Inwagen (1986, p. 207) admits he isn't) to point out where the argument goes wrong. That Lewis's argument is so far-reaching puts higher stakes on it, and therefore we may be inclined to doubt it more than we would otherwise, but it isn't, in and of itself, reason enough to move past it. We are, then, left with the claim that, if correct, Lewis's argument would be just as damaging for his own perspective, but just like in the case of Zaragoza's reply, here we seem to be again in the face of a *tu quoque*. If the argument succeeds, it is damaging for Lewis, but it does not exactly make Primitivist Ersatzism look any better.

[23] See van Inwagen (1986, p. 210) for an adaptation of Lewis's argument against the selection relation, replacing it with the membership relation.

As it turns out, Lewis bites the bullet. In his *Parts of Classes* (1991), Lewis is very critical of the notion of a singleton, which he very begrudgingly takes as a primitive (while in the process of writing the book, he came to learn of the possibility of doing away with it, via the notion of plural quantification, which he took to be good news, given his worries about how well understood the notion of a singleton was[24]). The notion of a class is usually introduced via such imagery as thinking of collections of objects and taking them together, thinking of them as a unity[25]. Common examples are for instance social groups, such as a board of administrators, a jury, wolf packs, a class of students. However, what is it we are talking about when we make a distinction between David Lewis and his corresponding singleton, {David Lewis}? In this case we are not given any multiplicity of objects we are uniting in thought or word: we have just one object. I agree with Lewis that the notion of a singleton is indeed obscure. And that seems enough to throw suspicion on how well we understand the relation of membership, for the singleton set of any object x is defined via the membership relation: {x} is the set of which x is a member and nothing else is. To these considerations, Lewis adds van Inwagen's (1986) case and the argument against Magical Ersatzism for yet another critique of the notion of a singleton: it seems that the set-theoretical membership relation would have to be an external relation that still, somehow, expressed a necessary connection between certain objects, which, as seen, is mysterious, given the intuitive strength of the recombination principle and the definition of what an external relation is.

Could we appeal here to the same strategy as we did in the case of the *having-a-length-in-meters-of* relation, and claim that perhaps set membership is an extrinsic relationship, so that it holds between members of sets and their respective sets in virtue of their relations to other objects? In contrast to the first relation, here it is not clear what objects would have to be related to both the member of the set and the set for there to be an extrinsic relation. Perhaps one could appeal to the properties contained in the axioms of a given set-theoretical system, or to the axioms themselves. An object would have to be a member of a given set, as a matter of necessity, because of the axioms stating how sets are to be constructed. That an object is always a member of its singleton would just be a *desideratum* that any plausible set-theoretical system would have to satisfy, but this relation would nonetheless depend on the specific way sets are taken to be constructed. Since these axioms would be in a certain sense prior to the sets (*i.e.* something is a set only if it satisfies the axioms), we would have then something akin to the case of the metre. Just as we only have a notion of a metre after we have fixed a given length and decide to consider it as a unit (relating it externally with the number one), we seem to only have a notion of a set in the context of an axiomatic theory (this doesn't imply, however, that we don't have any previous intuitive everyday notion of a collection, for instance) stating certain properties of it. This is all very speculative and I do not wish to rest my argument on it, so we move on to consider other hypotheses.

---

[24] See Lewis (1991, p. viii).

[25] See *Ibidem*, pp. 29-30.

Importantly for the resolution of this standoff is that there is available to Lewis a reply in the case of the membership relation that isn't available to the magical ersatzist in the case of the selection relation. For Lewis, to say that it is not possible that his cat, Possum, doesn't belong to {Possum} is to say that no counterpart of Possum does not belong to a counterpart of {Possum}. However, this is true by definition, the counterpart of {Possum} is defined in Lewis's perspective as the set which contains the counterpart of Possum and nothing else. So, even if the relation of Possum to its singleton is external, the necessary connection is still there as there is equally an external relation holding between every counterpart of Possum and a counterpart of {Possum}. This appeal to a counterpart-theoretical solution isn't available to the magical ersatzist, as the necessary connection between the total universe and the elements would hold in virtue of the intrinsic nature of the total universe (so we need to consider what elements its duplicates would select), not in virtue of the identity of the total universe (case in which we would have to consider its counterparts), like it seems to be the case for the relation between an individual and its singleton set.

We have, then, addressed both worries of van Inwagen's: on the one hand, perhaps it is right after all that the membership relation isn't well understood; on the other, it seems that Lewis has a way of, in his ontology, explaining the necessary connection between an individual and its singleton. One important upshot from this debate seems to be that if one does not accept Lewis's ontology, as I and impossible world theorists don't (or at least they can't accept that it tells the full story, if they accept an hybrid account (see Berto (2010) and Berto and Jago (2019, pp.58-61))), is that there seems to be an open question whether the notion of membership is well understood, and how to resolve the conundrum.

In the last few pages we have considered different ways proposed so far of trying to hold that a selection relation between the total universe and abstract simples is external, and we found them lacking. Some of them consist in a *tu quoque* and in that regard are not totally satisfying, especially as here our purpose is not to defend Lewis's modal realism from its critics, but rather assess the merits of Magical Ersatzism. Still, ignoring this last fact, we found out that there are reasons to think that the arguments considered are not decisive replies to Lewis's criticism of taking the selection relation to be external. We move on to consider the possibility of it being an internal relation.

Dealing with the dilemma: the relation is internal

One important point that Lewis (1991, p. 35) notes in regards to the membership relation that will allow us to move the discussion further on whether the selection relation might be taken to be internal goes as follows: "it's no good saying that a singleton has x as its member because it has the property: being the singleton of x. That's just to go in a circle. We've named a property; but all we know about the property that bears this name is that it's the property, we know not what, that distinguishes the singleton of x from all other singletons.". This particular point displays, as Fisher (2015) stresses, an important characteristic of Lewis's methodology, a characteristic that one needs to keep in mind if one is to understand the objection to taking the selection relation to be internal.

As van Inwagen (1986) points out, “possible world” is in several contexts a functional term, that is, an expression that denotes whatever it is that fulfils a certain role. When applying it, then, Lewis and ersatzists can have fruitful conversations about possible worlds without talking past each other and may even agree in a lot of respects about how to put them to work, giving support to each other’s perspectives. A clear example is the so-called Lewis-Stalnaker analysis of counterfactuals. So, when philosophers argue about what possible worlds *are*, they are arguing about what entities, if any, should fulfil the *ways-the-world-can-be* role. Lewis (1986, p. 140) explicitly endorses such a description, and the point that should be stressed in regards to the present discussion follows directly from it: in the context of a debate about what entities fulfil a given role, it does not help at all to be told that they are the entities that fulfil the role[26]. It wouldn’t, for instance, be clarificatory, in the context of a metaphysical debate about what the nature of possible worlds is, to be told that they just are ways that the world can be. And in just this consists the feature of Lewis’s methodology that matters to us at this point: that one doesn’t really have a notion of what a concept X stands for if the only thing one can say about X is that it fulfils the role it is introduced to fulfil; and if one doesn’t have that notion, then we don’t really grasp the concept we have introduced and that we thought we understood.

Denby’s (2006) criticism of Lewis’s argument against the selection relation being internal, Fisher (2015) claims[27], displays a failure to appreciate the import of Lewis’s methodology for the broader argument against Magical Ersatzism. For instance, Denby’s (2006, pp. 166-168) criticism of Lewis only focuses on the absence of acquaintance with most terms in the pairing of the total universe’s intrinsic properties with the elements’ intrinsic properties. And here I agree with Fisher (2015) that Denby misses the point[28], or at least one of the most important points of Lewis’s argument: that we don’t actually have any grasp of what the intrinsic

[26] I believe that it would even be straight out infelicitous to say it. To see why, notice that it is rather common when speakers say sentences of the form “The x is the x” that the listeners don’t take them to be speaking literally, but rather to be, for instance, jesting. When I ask a friend who the last king of France was, they may reply in a joking tone “the last king of France was the last king of France”, case in which I understand him to be telling me that he doesn’t know who it is, or to not be giving a serious answer. The situation here is similar, but just so happens that since the expressions used (for instance “ways the world can be” and “possible worlds”) are different, there might be an illusion of something informative having been conveyed, *in the context of this specific debate*. If one is presenting possible world semantics to someone that is wholly unfamiliar with it, saying that they are ways that the world can be would in some cases be informative. But perhaps what would be informative in such contexts would be that the listener would then come to know what the functional role at hand is (and one can say a lot about the role without specifying what fulfils it).
[27] Fisher (2015) also mentions that Zaragoza (2007) fails to convey this important information, and while true, this seems unfair: Zaragoza agrees with Lewis that the relation of selection cannot be internal, and spends most of his paper trying to prove that it can be external. For his purposes, Zaragoza doesn’t need to provide any further information on this part of Lewis’s methodology, as it doesn’t play any significant role in the discussion of whether the selection relation could be external.
[28] Here I won’t be going over Denby’s argument. See Fisher (2015, p. 11, n. 6) for a rebuttal of their argument that relies on the present considerations, and that I take to be fully right.

properties of the elements would be, for the same reason we would not have a grasp on the concept of a possible world if we could only say about it that a possible world is a way that the world can be. Lewis seems to argue for that conclusion in the following way: given that we have an element for every difference in the state of the world, we need an infinite number of distinct intrinsic natures of the elements; but, given their simplicity and them being causally disconnected from us, what could these natures be?; the only plausible candidates seem to be properties of the form *represents that so and so* where *so and so* just is a certain state of the total universe or of a part of it; but here the use of the expression "represents" is question-begging, for it is precisely to account for worldly representation (even if not by providing an analysis) that we have introduced the abstract simples and the selection relation: to say that the elements that represent that *a* is *F* are selected partially in virtue of their intrinsic natures, which turn out to be that they are the ones that represent that *a* is *F* (together with what differences there are between them) is plainly and viciously circular; therefore, we don't really grasp what the intrinsic natures of the elements are.

There is a further argument presented by Denby (2006, pp. 168-170) that, while also stemming from a misunderstanding about the point of Lewis's overall line of argumentation, will allow us to see Lewis's methodology in action, as well as to provide a bridge for the discussion that follows. The argument is that "selects" might just be a vague term, just as many predicates of natural language are. The magical ersatzist would just have to add that we are also unable to grasp any precisification of "selects". We wouldn't, then, have to be acquainted with any of the terms of the pairing between the abstract elements' natures and the states of the total universe.

I believe the argument is misguided from the get-go, as we are dealing here with a very strong notion of "grasping" that seems to be suited only for inquiries as the one we are occupied with. When Lewis complains that if we just say that possible worlds are ways that the world can be we haven't gotten a grasp of the notion of a possible world, this complaint seems to only be plausible if we assume a rather strong notion of what it is to grasp a concept. Students taking a first course in modal logic seem to be able to use effectively, in some ways, the notion of a possible world, even if they may have a very imperfect understanding of it. It must not be such a weak notion of graspability that is at hand. In fact, Lewis's insatisfaction seems to mirror a trend in philosophical exploration of certain themes that helps explain how complex but detailed reasoning came to be one of the standards of contemporary analytic Philosophy. In various instances throughout the history of Philosophy, starting with the Socratic dialogues, arguments have been presented for why we don't actually grasp adequately certain concepts we thought we did. For instance, since he was not able to tell what piety is, Euthyphro was concluded to not actually have a clear understanding of the concept he professed to have a clear grasp on[29]. It is doubtful, however, that for the purposes of daily life Euthyphro wasn't able to use "piety" appropriately. Lewis's considerations in certain respects resemble the Socratic method in this way. Can we say that we grasp in this strong sense a vague predicate? It seems that in some cases

[29] Plato (2005, pp. 1-59).

yes: it seems that to have the concept of baldness, we don't have to (and *must not*) present a precise definition that makes a clear distinction between bald and non-bald people. But are we in the same position when it comes to vague predicates we can't in principle grasp any precisifications of? In the case of baldness, it seems that we can understand stipulations that change the concept, so it corresponds with one of its precisifications, for instance if "bald" started to refer only to people with 500 hairs on their scalps or less. In the case of "selects", however, we are unable to grasp any of its precisifications. Denby (2006) correctly claims that someone with no understanding of what water-droplets are can still grasp what a cloud is. But is the understanding of such an agent enough to count as grasping in this stronger sense? It seems that that's not the case, for if they had enough hold of the concept, they would be able to tell what clouds *are*, and that, not knowing anything about water-droplets, the agent seems to be unable to do.

This point is important, as it makes more plausible the earlier claim that perhaps the set-theoretical notion of membership hasn't been adequately grasped after all. Van Inwagen (1986) worries that the conclusion of Lewis's argument leads us to the conclusion that classic mathematics is confused, ungraspable. This could evoke fears that whole branches of mathematics are to be considered as nothing short of incomprehensible gibberish. But that is not what is being claimed. What Lewis's argument establishes is merely that in this strong sense we don't grasp the fundamental notion we have resorted to. Even with a very imperfect grasp on it, however, we can still make a series of claims taking for granted that the concept has been understood. It all seems to come down to what we are using the concepts for. As noted earlier, Lewis (1991) himself begrudgingly accepts the notion of a singleton as a primitive. If his stance on what isn't (in this strong sense) graspable was so extreme as to count it all as nonsensical, it would have been incoherent to accept it for whatever purposes. Instead, we should suit our requirements for grasping to the context at hand. Perhaps for doing most of what mathematicians are occupied with, the standard notion of what a singleton is may be enough, but when the issue at hand is the nature of sets and the foundations of set theory, the standard notion seems to prove lacking. Similarly, with "cloud" as used in everyday life and in a Physics class; and with "selection" in some theoretical uses for possible world metaphysics (if we can grasp the concept at all) and in the context of discussing how possible worlds represent what they do. As van Inwagen (1986) puts it, for Lewis it's a matter of the theory "doing the work that it's supposed to do", and in the present context the "work" is providing an account of worldly representation.

One other problem that this argument faces was already noted right at the beginning of the present discussion of Magical Ersatzism, and ends up tying in neatly with what has just been said. The strategy of stating that it does not matter what the precise selection relation is like, as long as the implications between the elements and the relational system remains the same, seems to only work in the context of doing possible world semantics, in which we can take worlds to be whatever objects we may like. As noted, it is also not convincing, in the context of this reply, that we need new abstract simple objects to do the job at hand: why not have, for instance, the pure sets stand for the possible worlds? Why introduce a new kind of abstract, simple objects? Let us set these last worries aside for now. It was noted in the beginning of the present treatment of Magical Ersatzism that the reason this approach fails comes down to the fact that, in the context of the debate of what possible worlds are, there seemingly being a need to say more about their nature. This seems to be, in a way, the same worry that Lewis would have with this proposal: the Magical Ersatzist hasn't said enough (and Dunbey (2006, pp. 168-9) on this proposal accepts that they can't) about what possible worlds are, so that we don't have a theory capable of doing the work that it is supposed to do.

Doing away with the mystery by specification?

There is a last argument that, given the whole discussion so far, is easy to see where it goes wrong. It starts off from a grudge against the use of the term "selection". When we describe Magical Ersatzism in such general terms such as "selection" and "abstract simples", the mystery ensues. But, it is claimed, this mystery is just apparent and is an artifact of the way the position is being presented. If instead of "element" we say "proposition", or "state of affairs" or "property", and instead of "selects" we say, respectively, "makes-true", "obtains" or "instantiates", the mystery evaporates. Everyone, they continue, understands what making a proposition true consists in, or what it is for a state of affairs to obtain, or what it is for an object to instantiate certain properties. And so, after all, we seem to have a perfectly good handle on these relations after all.

Lewis (1986) replies that simply putting in "proposition", say, for "element" won't help with the problem: we have just switched the terminology. In this sense of grasping that we are assuming, it is then not clear that we do grasp what "proposition" means in this context. It seems merely like a different placeholder name for what one would wish to precede a further elaboration of what possible worlds and propositions are. It might be said in reply: but propositions are also meant to play other roles, so it is informative after all to change the terminology, namely, we are saying that what fulfils certain other roles also fulfils the *ways-the-world-can-be* role. A follow-up worry would be, of course, that this just pushes the mystery further, for even if it is right that a concept such as "proposition" is supposed to serve a number of different roles, it still is, nonetheless, a functional concept. What we need, again resorting to van Inwagen's (1986, p. 192) distinction, is an ontological concept that in this discussion would tell us more about what fulfils the specified roles. While saying that "proposition" fulfils the ways the world can be role requires further restrictions in what kinds of objects can fulfil the multiple roles united under the functional concept of a proposition, it does not yet tell us what a proposition - and therefore a possible world - is.

It seems that the only way to get out of this conundrum would have to be to present independent arguments for the existence of propositions, states of affairs, or properties, and to say more about their nature. Recall, however, that these are all being taken to be abstract and simple. As Jubien (1991) ends up arguing (even if in a misguided way about the meaning of his conclusions (see Fisher (2015))), it seems that only assuming that any of these entities is complex can one argue for their existence and say more about what their nature is. Not assuming any structure, it seems that the only path available would be to try to specify their nature in relation to the role they're supposed to play, but that, as we have been insisting on, seems untenable.

One important conclusion that has come out of the last paragraphs is that it is only because of the *simplicity* of the elements that it seems we are unable to resort to independent explanations and, at various points, to a more satisfying account of what possible worlds are. We proceed, then, to consider proposals according to which possible worlds are abstract constructions of some kind. These will all be considered under a very broad understanding of "Linguistic Ersatzism", which we may call Constructivist Ersatzism.

Constructivist Ersatzism

Having taken note of the challenges to the primitivist view, we now turn to consider linguistic ersatzism, understood in a very broad sense, or what we might call constructivist ersatzism, that is, the family of views which takes worlds to represent, for instance, that *a* is *F* without thereby containing *a* as a part, while also taking worlds to be abstract constructions of some kind. As we saw in the last section, the prospects for both explaining worldly representation whilst taking worlds to be abstract simples, and refusing to give any explanation of said representation aren't very promising. We now focus on what seems to be the remaining option for ersatzists: to claim that worlds are abstract constructions out of simples.

Many views can be grouped under this general description, such as: that worlds are maximally consistent sets of propositions; that worlds are maximal and consistent states of affairs; and that worlds are maximally consistent stories in a world-making language (that is, sets of sentences). The only requirement is that whatever worlds are taken to be, they are taken to be complex.

In this section, as in the last one, we start by giving a brief characterization of the views under discussion, and then move on to consider some of the main arguments against them. Since claiming that a world represents that such-and-such is a primitive to not be further explained led us to maintain a poorly understood selection relation, we must then explain how it is that worlds represent what they do. Given that the characterization above of linguistic ersatzism is very general, here some of the most influential versions of constructivist ersatzism will be briefly presented, alongside questions raised for them in particular but which do not affect other versions of linguistic ersatzism. Finally, general objections against constructivism will be considered.

The most serious problems for all views that purport to explain how it is that worlds represent while adhering to constructivist ersatzism arise when we consider the question of what resources are available in the actual world to construct the kinds of worlds the constructivist appeals to, and if they are enough to make all the distinctions between possibilities one would want there to be. In particular, statements about non-actual objects seem to be impossible to assess from an actualist perspective, as do statements containing non-actual but possible fundamental properties. These objections will be considered later on.

Possible worlds as maximal states of affairs

The first form of constructivist ersatzism here explored takes worlds to be maximal states of affairs, that is, states of affairs such that they either "include or preclude"[30] every other state of

[30] Berto and Jago (2019, p.74).

affairs. These notions of inclusion and preclusion, as Berto and Jago (2019, pp. 74-75) point out, are inherently modal: a given state of affairs, *s*, precludes another, *s'*, if and only if it is not possible for *s* and *s'* to obtain together; and a given state of affairs, *s*, includes another, *s'*, if and only if it is not possible for *s* to obtain without *s'* obtaining. In so far as this characterization is itself modal, ersatzists that subscribe to this view give up on the project of providing a reductionist account of modality.

Unfortunately, the notion of a maximal state of affairs as so far presented will not work. To see why, consider the fact that the state of affairs corresponding to A & ~A (*something's being A and not being A*), would thereby preclude every other state of affairs (for, given that it is not possible for this state of affairs to obtain, it is not possible for it to obtain together with any other state of affairs). Since it precludes every state of affairs, this contradictory state of affairs would then be maximal. But then we can no longer identify the possible worlds with those states of affairs that are maximal.

Another way of understanding this notion of a possible world - but that also appeals implicitly to modal terms - is to consider maximal states of affairs to be instead mereological sums of states of affairs, such that for any given pair of incompatible states of affairs, *s'* and *s''*, that is, states of affairs that preclude one another, either one or the other is included as a part in a given maximal state of affairs, *s*, but not both[31]. Whereas in the first alternative, the implicit appeal to modal terms is present in the use of the notions of inclusion and preclusion between states of affairs, here primitive modality is present in the notion of two incompatible states of affairs. A further worry for this proposal would be that a pair of contradictory states of affairs would be individually incompatible, and therefore could not obtain by themselves or together, and so we would have that one of them would have to be included as a part in a maximal state of affairs. But that would lead to one of two unwelcome conclusions: either that i) contradictions are possibly true; or that ii) possible worlds are not, after all, to be identified with maximal states of affairs. One way to remedy this is simply to restrict the states of affairs to the possible ones, that is, to the ones that possibly obtain, so we would have that: a given state of affairs *s* is maximal/a possible world if and only if for any given pair of possible but incompatible states of affairs, *s'* and *s''*, either *s'* or *s''* is part of *s*. Here we would resort yet again to primitive modality, but as mentioned earlier, providing a reductive account of modality is a feature that the states of affairs approach does without by default.

Crucial to the view of possible worlds as maximal states of affairs is the notion of an existing but non-obtaining state of affairs. On this view, what distinguishes the actual world from the non-actual possible worlds is that whereas the maximal state of affairs that corresponds to the actual world both exists and obtains, all the other maximal states of affairs exist but do not obtain. Insofar as this view is tenable, then, one must be clear about what the notion of a state of affairs obtaining comes down to. Berto and Jago (2019, pp. 76-77) argue that the notion of a non-obtaining state of affairs is hard to make sense of, and reject the view of taking worlds to be

[31] Berto and Jago (2019, pp. 75-76).

states of affairs on that basis. Here we may remain neutral on this question, for there seem to be independent reasons to reject taking possible worlds to be maximal states of affairs.

The notion of a state of affairs has been subject to lengthy discussion, and a good deal of suspicion. One intuitive notion seems to run as follows: when we truly say, for instance, that "A cat is on the mat", our sentence would express the state of affairs of *a cat's being on the mat*, that is, a complex entity which includes the cat, the mat, and the relation holding between them[32]. This complex structure, given that the sentence is supposed to be true, is said to obtain, whereas false sentences would represent states of affairs that do not obtain. Here the notion of a state of affairs obtaining might for instance be elucidated by the notion of truth: a given state of affairs obtains if and only if the proposition that represents its content is true.

Views that take states of affairs to be complex have been met with strong criticism, however. A first objection to taking states of affairs to be complex is that it seems we run into contradiction if we claim that the particulars, properties and relations involved in the states of affairs are parts of those very states of affairs, namely, since the relation of parthood is transitive, we would have that parts of parts of states of affairs are themselves parts of states of affairs. For instance, the state of affairs *Vesuvius's being a volcano* would, therefore, contain Vesuvius's lava, whereas it seems that the state of affairs in question does not contain any lava (as Frege writes to Wittgenstein in a letter from 1919[33]). A second objection is that it seems that a given state of affairs can exist even if one of the particulars involved does not: *a's beginning to exist when b ceases to exist* seems to be a legitimate state of affairs, but, given the relation expressed between *a* and *b*, the state of affairs can only obtain if *a* begins to exist but *b* does not exist anymore[34]. It is a mystery, then, how this state of affairs could obtain when one of its parts does not exist. A third objection is that according to classical mereology, wholes built up from the exact same parts are equal, and yet the states of affairs *Othello's loving Desdemona* and *Desdemona's loving Othello* are distinct, despite containing as parts the same particulars and relation. A fourth objection is that if we take states of affairs to have parts, and these parts to be the individuals, properties and relations that figure in the states of affairs, then we will reach cases in which a state of affairs has only one proper part, contrary to the mereological principle of the remainder, or the principle of weak supplementation[35].

---

[32] Alternatively, some philosophers, such as Wittgenstein (1921), think of states of affairs as being complexes formed only of objects which by themselves fit directly with one another, case in which there is no need for an over and above constituent of the state of affairs that we may call the relation that holds between the objects featured in it. What is said here about the state of affairs view does not, in any way, rely on the specifics of either conception of state of affairs. In the above characterization, then, "relation" can be read both as a further term binding the objects together, or as a particular way in which the objects by themselves are arranged.

[33] Textor (2021).

[34] See Künne (2003, p. 122).

[35] This principle holds that whenever x is a proper part of y, there is a z that is part of y and which does not overlap (that is, has no part in common) with x. Informally we can say that if an object has a part that is not identical to itself, then if we remove said part, *something* of the initial object will remain.

Finally, there are some worries about the unity of states of affairs. What, if anything, binds the constituents of the states of affairs together so that we can take each one of them to be *one* state of affairs? It seems that it cannot be anything that is external to the state of affairs, in the sense of being something other than the particulars and properties involved, for otherwise Bradley's regress would ensue, as we would then have to explain what is it that binds the constituents plus this new binding item all together, and then what binds this last binding item with the other constituents, and so on. Whatever it is that unifies the constituents of a state of affairs cannot, then, be itself a further constituent of the state of affairs. It seems, then, that nothing else except the constituents of the state of affairs unify it. One way to do so is to maintain that a particular, *a*, and a property, *F*, are unified into a state of affairs whenever *F* is predicable of *a* (Johnston (2006, p. 684), which goes in line with Berto and Jago's (2019, p. 76) claim that states of affairs are introduced to make sense of the notion of predication).

Here there is a further question, however, about whether whenever a property is predicable of a particular there is a corresponding state of affairs, or whether we should claim that only in some cases is there a state of affairs. Textor (2021) argues for the latter position, but here we need not take a stance on this issue.

Another route one might take here is to follow Fine (1982) and consider states of affairs to be themselves simple but ontologically dependent on the objects we first thought to be the constituents of the states of affairs. So for instance while the state of affairs of *a cat's being on the mat* is itself simple, it depends, for its existence, on the further existence of the cat and the mat (and perhaps the relation *being on*). As Textor (2021) goes on to show, this position would be able to meet the objections so far considered (for instance, the problem of the unity of a state of affairs dissolves as states of affairs don't have constituents). I believe, however, that this way out faces the same objection that Lewis has raised for magical ersatzism, for if we say that states of affairs are abstract simples, but that nonetheless are ontologically dependent on the particulars and relations that we first thought were their constituents, then we're faced with the question of why it is *that* simple abstract object dependent on the particulars and relation in question and not some other abstract simple - that is, of what allows us to pick one simple instead of any other to play the role of *that* state of affairs instead of any other. Here the objection would have to be met as we are, again, trying to give an account of the *nature* of possible worlds, so that our discussion is ontological, not semantical.

Mirroring the discussion from the last section, we might then start questioning whether the relation of ontological dependence is internal or external. If it is external, then it seems that any other abstract simple would do the job. This, however, seems to go against what we intuitively take ontological dependence to be, in the sense that it seems that a given object, *O*, is ontologically dependent on some other object, *o*, only if the existence of *O* necessitates the existence of *o* (that is, *O* cannot exist without *o* also existing)[36]. But if any element will do, then

[36] Necessary coexistence is only a necessary condition for ontological dependence. Suppose that the empty set is an abstract object and that it necessarily exists, then for whatever object, it cannot exist unless the empty set exists, for the empty set exists in all possible worlds. We wouldn't, on this basis, claim that any object whatsoever depends for its existence on the existence of the empty set.

it seems that we can have a given state of affairs, *s*, which ontologically depends on certain particulars and relations, correspond to the element *e1*, but could have very well had *s* correspond to *e2* instead, with *e1* corresponding instead to a different state of affairs *s'*. If this were so, however, then in some cases the objects on which the state of affairs *s'* depends on could exist without the objects on which the state of affairs *s* depends existing. This would be a case, then, where *e1* would exist (since in this case it corresponds to the state of affairs *s'*, not *s*), but not the objects on which *s* (the state of affairs that we first took *e1* to correspond to) ontologically depends. But that means that *e1* does not ontologically depend on the objects that the state of affairs we initially took it to be depends on. The relation of ontological dependence, thereby, cannot be external.

Bar taking the relation to be extrinsic, which is not plausible for similar reasons to the case of magical ersatzism (Lewis (1986, p. 182)), it seems that the only option left is to take the relation of ontological dependence to be internal. But here again we can mirror Lewis's discussion of magical ersatzism. It seems that if we imagine a continuum of spacetime points and we take every different configuration of matter on those points to be possible, then it seems we will have for each spacetime point, *st*, the corresponding states of affairs *st being occupied by matter* and *st not being occupied by matter*. But if this is the case, then we will have infinitely many simple states of affairs (that is, states of affairs that do not contain as parts any other state of affairs but themselves). Given that the elements are simples, it seems that their intrinsic properties will have to be given in terms of an infinite pairing between all the states of affairs and the objects and relations they ontologically depend on. Just like in the case of magical ersatzism, it doesn't seem here that we can truly grasp such an infinite pairing. If we understand why a given state of affairs ontologically depends on given objects and relations, it seems to be only resorting to some kind of magic.

The point of the preceding paragraph generalizes. In the section relative to magical ersatzism we saw that Lewis's objection against taking worlds to be abstract simples that are in some sense maximal holds because of the lack of structure of the abstract simples. If the reasoning in the preceding paragraph is right, then we can expand that objection to some sorts of constructivist ersatzism: if we take worlds to be constructions out of abstract simples, then there will still be a question of why it is certain abstract simples that are playing the specific roles that they play in our theory, and not any other (for instance, why can't the abstract simples that correspond to the states of affairs *s* and *s'* exchange roles?).

It seems, then, that we should opt for a kind of constructivism according to which the possible worlds are not constructions out of abstract simples. The objections seen against taking states of affairs to be structured entities, however, still stand. It remains to be seen whether these objections can be met. Later on we will come to some of these difficulties. Here, however, is not the place for a full assessment of the coherence of the notion of states of affairs, which should include both a discussion of the objections just mentioned, as well as an elucidation of the seemingly mysterious notion of an existing but non-obtaining state of affairs. Pending replies to the objections so far considered for a theory of possible worlds as states of affairs, and in particular, pending an elucidation of what it is for a state of affairs to exist but not obtain, it would seem that the linguistic ersatzist should, for now, look elsewhere for an account of the nature of possible worlds.

Possible worlds as maximally consistent sets of propositions

As we have just seen, whatever possible worlds are taken to be on the linguistic esrsatzism approach, they should not be taken to be constructions out of abstract simples, for if they are taken to be so, Lewis's (1986) objection against magical ersatzsim would also be applicable to linguistic ersatzism, and therefore this latter approach would lose its edge over the former.

Whatever propositions are taken to be, however, it seems that they are abstract: propositions do not seem to be located in any spacetime-like structure[37]. It seems, then, that on the proposal of worlds as maximally consistent sets of propositions, propositions can't, therefore, be taken as themselves being simple, and must be structured entities. One way to spell this out would be to take propositions to be constituted of Fregean senses, such that the proposition expressed by a given sentence would be a function of the senses of the constituents of the sentence. Fregean senses, however, are themselves paradigmatic examples of abstract objects, constituting what Frege (1956, p. 302) refered to as the third realm, on the side of the mental and physical. On this view, worlds would be sets of propositions, with propositions themselves being structured entities formed out of other abstract objects, the Fregean senses.

If we take this Fregean route, there is a further question about whether Fregean senses are themselves simple or complex. If they are simple, it seems that a new variation of Lewis's (1986) objection rears its head: why is it a given abstract simple that serves the role of being a mode of presentation of *this* referent instead of any other abstract simple? Just like in the case of possible worlds and what they represent and in the case of relations of ontological dependence, it seems that also here there is a necessary connection between the objects referred by a given term, for instance the moon, and its modes of presentation, that is, the corresponding Fregean senses, for if there was no moon to begin with, then it seems that there would not exist any mode of presentation of the moon either. And that's all we need to be able to raise Lewis's dilemma again.

We seem to have, then, reached a further conclusion about what a linguistic ersatzism approach to the nature of possible worlds should look like: taking worlds to be abstract constructions of some kind, still, whatever their ultimate constituents are (that is, the constituents of worlds that do not have constituents themselves), they cannot be abstract simples. But if they are not abstract simples, what else could they be? Abstract complex objects seem to be ruled out, for in that case they would not be ultimate constituents, as they would themselves have

[37] Here it could be claimed that the proposition that *Socrates is a philosopher* is wherever Socrates instantiates that property, which in turn would presumably be wherever Socrates is. This, however, would not work. On presentist views, Socrates does not exist anymore, so there is no place where Socrates is, and therefore no place where he instantiates the relevant property. Yet, there does seem to exist such a proposition, and indeed that it is true. A further worry for such a view is that all false propositions seem to have just as good a claim to existence as the true ones, and yet it is not clear where propositions like *The moon is made of green cheese* would be located.

constituents. Simple and complex *abstracta*, however, seem to correspond to all abstract objects. We are, then, led to conclude that the ultimate constituents of worlds should be concrete, not abstract[38].

Coming back to the linguistic ersatzism view at hand, it seems that propositions cannot, therefore, be either abstract simples or be made out of abstract simples. We are then left with the option of taking propositions to be complex abstract objects made out of concrete objects. One way to do so is to take propositions to be Russellian tuples, that is, ordered tuples whose constituents are certain particulars and universals (given present worries, perhaps the easiest way of making the theory run would be to consider universals to be immanent, and therefore concrete). So for instance the proposition that *Othello loves Desdemona* would be the ordered triple <Othello, *loving*, Desdemona>[39].

Thinking of possible worlds in terms of maximally consistent sets of propositions instead of propositions as sets of possible worlds has, as for instance Stalnaker (1976) has pointed out, the consequence of taking away one of the main reasons for accepting possible worlds in one's ontology: to make sense of the notion of proposition. Further, it has been argued, for instance by Jago (2015), that understanding propositions in this way leads to a failure of accounting for some of the essential features of propositions. Jago (2015, p. 599) claims that two essential features of propositions are that they are true and false at worlds, as well as possible, necessary or impossible. Accepting Fine's (1994) considerations on essence, Jago endorses what he refers to as the *Nature of Sets Thesis*.

Nature of Sets Thesis: If X is a member or subset of Y, then it is of the nature of Y that X is one of its members or subsets, but not of X's nature that it is a member or subset of Y.

This thesis can perhaps sound strange at first glance, but some examples help make it sound more plausible. Arguing against modal accounts of essence, Fine (1994, pp. 4-5) gives Socrates and his corresponding singleton as examples of what Jago (2015) takes to be a more general point about the nature of sets. While in all possible worlds in which Socrates exists, so does his singleton, it does not seem right to claim that it is part of Socrates's essence to be the sole member of his singleton. Presumably, Socrates's essence has something to do with him being a member of the species *homo sapiens sapiens*, or with him being an embodied mind, or whatever else one takes persons to essentially be (assuming, as sounds intuitively plausible, that Socrates is essentially a person). So, argues Fine, what is part of Socrates's essence is not the same as what is true of Socrates in all possible worlds in which Socrates exists. Here are two further examples. Everything is necessarily self-identical, yet it does not seem that the property of being self-identical is part of every object's essence. If we allow for complex properties, then we can say that *being such that 2 + 2 = 4* is a property that Socrates has in all possible worlds

[38] If one takes the constituents of propositions to be particulars and universals in the case of singular propositions, or universals in the case of general propositions, perhaps a further argument against transcendent universals - that is, abstract universals with no spatiotemporal location - and in favour of immanent universals - that is, universals wholly located in all of their instances - could be extracted from the present considerations about structure.

[39] See King (1995) for an elaboration of the view of propositions as Russellian tuples.

(the same applies if we replace *2 + 2 = 4* by any other necessary truth), yet likewise it seems wrong to claim that *being such that 2 + 2 = 4* is part of Socrates's essence. If these examples have convinced you, then you're on board with Fine's critique of the notion of essence as what properties an object has in all possible worlds in which it exists. Of course, even accepting Fine's argument, that an object has certain properties in all possible worlds in which it exists can still be considered as a necessary condition for those properties to constitute its essence - the important aspect is that if Fine is correct, this cannot amount to the full story, that is, it cannot be a sufficient condition for certain properties to constitute an object's essence.

Going back to the first example, Jago (2015) takes Fine's (1994) line of argument to show a more general feature pertaining to the nature of sets, the aforementioned *Nature of Sets Thesis*. Let us now apply the *Nature of Sets Thesis* to the question of what propositions and possible worlds are. If possible worlds are maximally consistent sets of propositions, then it will be of a given world's nature that it contains certain propositions, and not of a given proposition's nature to be part of any specific possible worlds. In a possible worlds framework, for a proposition, *p*, to be true at a given possible world, *w*, just is for *w* to be a member of *p* (if propositions are taken to be sets of possible worlds), or for *p* to be a member of *w* (if possible worlds are taken to be maximally consistent sets of propositions), so to say that it is of the nature of *w* to have *p* as a member is the same as saying that it is of *w*'s nature that *p* is true at *w*, and not the other way around. However, as we just saw, it is of propositions' natures that they are true at worlds, so this picture of possible worlds as maximally consistent sets of propositions, which are in turn taken to be Russellian tuples, doesn't seem to be able to account for one of the essential features of propositions[40].

Jago's argument is interesting and has a lot going for it. Yet, I believe that a friend of a structured notion of propositions can give the following reply to it. While it is certainly the case that propositions are, by their very nature, true or false, and possible, impossible or necessary, still, it is not of a proposition's, *p*, nature alone (and so, not of its essence) that it is true at a specific world *w*, but rather of the nature of the proposition *and* the world it is true at. *Being-true-at* might be taken to be an internal relation holding between propositions and worlds, and if it holds between *p* and *w*, then it will be the case that *p* has the property *being-true-at-w*, but then if *p* has the relevant property or not will not depend solely on its nature. Perhaps similarly to be an individual may just be to possess certain properties, but it is only in function of the nature of the properties together with the nature of the individual that one can determine what properties an individual has (contingently or necessarily). Another example is that perhaps to be a living being is to belong to some species or another, but to which species a given living being belongs to depends not only on the properties of the individual, but also on the properties of the species, so that to be a member of the species *homo sapiens sapiens* is not just to be an organism with certain characteristics, but also for those characteristics to be shared by other organisms in a way such that they can be grouped under a same species. We can put it differently by saying that it is of propositions' nature to be true or false at worlds but it is not of their nature alone what worlds

[40] A very similar argument is given in Jago (2015, p. 599) also for why this proposal fails to account for the essential feature of propositions of being possible/necessary/impossible.

in specific they are true/false at. I believe that a reply following these or similar lines could be successful, and, in so far as one accepts both the *Nature of Sets Thesis* and that propositions are sets of possible worlds, it could even raise a challenge to the opposing view (as it would have the wrong consequence that it is of a proposition's essence alone what specific worlds it is true/false at). Here this hypothesis will not be developed, and we can concede that Jago is right, looking elsewhere for a way of constructing the possible worlds of the linguistic ersatzist perspective.

Possible worlds as maximally consistent sets of sentences

As pointed out above, we're considering linguistic ersatzist positions, understood in the very broad sense of worlds being abstract constructions of some kind. Now we go on to consider a position that we might call proper linguistic ersatzism, which maintains that possible worlds are maximally consistent sets of sentences in a suitable worldmaking language. On this view, possible worlds are taken to represent just as sentences in a language do, namely, by containing the relevant worldmaking sentences. The nature of the worldmaking language is especially relevant, as the plausibility of the proposal under consideration will turn on whether circularity and expressive power worries can be met.

We can start delimiting the list of candidates to play the role of a worldmaking language by excluding natural languages from consideration. One of the aims of possible worlds semantics is to account for regular modal talk. Let us suppose, for instance, that we want to explain why it is true to say the English sentence "Donald Trump lost the election, but he could have won" (let us call it S). Let us suppose that possible worlds were taken to be maximally consistent sets of English sentences. Then it would be true to say S because the actual world contains S. But such an explanation does not seem satisfactory, for if we don't know why the original sentence is true, we won't come to know why via the explanation given in terms of the actual world containing S, as what distinguishes the actual world from all the other possible worlds, on this view, is that all and only the sentences contained in the actual world are true, so that the explanation would look like: S is true because it is contained in the set of sentences containing all true sentences. But if we don't know why S is true, it seems that we equally don't know why S belongs to the maximally consistent set of *true* sentences. Now suppose that worlds were given as maximally consistent sets of sentences in a different natural language, Portuguese, for instance. Then the explanation would look like: S, the English sentence, is true because S*, the translation of S into Portuguese, belongs to the maximally consistent set of sentences in Portuguese whose members are all true. This also does not seem to help, for it is not at all clear why we should pick one natural language instead of any other to function as our worldmaking language - it seems arbitrary to account for the truth of certain modal claims in any given natural language in terms of what modal claims are true in any other natural language. We have then reached a dilemma for taking the worldmaking language to be a natural language: either we explain what makes certain modal claims in a given natural language true in terms of possible worlds constructed out of sentences of that very language, case in which we seem to get a circular explanation; or we explain the truth of those modal claims in a different natural language, case in which the move seems *ad hoc*, to go against the notion that natural languages should be on a par, with no one language being privileged in terms of portraying the modal facts, and, in the end, to equally not provide a good explanation for why certain modal claims are true. We should, then, look elsewhere in order to find a worldmaking language.

A second worry pertains to the number of possibilities that can be represented with sets of sentences alone. In his book *Counterfactuals*, Lewis (1973) claims that we should allow for at least as many possibilities as the cardinality of the power set of the continuum. To see why, consider the example given earlier of a continuum of spatiotemporal points, each of which may or may not be occupied by matter. If we agree with Lewis (1973, p. 89) that to every different distribution of matter over the points there should correspond a different fully determinate possibility, then it seems we would need a different possible world for each such distribution of matter. Given that there would be continuum-many points, the amount of possible configurations of matter over the points seems to be equal to the cardinality of the power set of the continuum set. However, it seems that languages that can be learned by a human-like agent are finite, that is, they contain finite strings of symbols, which are themselves also finite (Bricker (1987)), so that they contain at most countably many sentences, and there wouldn't, therefore, be enough sentences to make all the distinctions between possible worlds we want to make. This approach, then, would seem to fail in virtue of there not being a language with enough resources to express all the possibilities without conflating them, whereas we want to explain what all the different possibilities are. We have failed, then, to provide a rich enough ontological story for what the nature of all the possibilities are.

This latter objection, which we may call the cardinality objection to proper linguistic ersatzism, presupposes that the worldmaking language at hand is finitary. This seems to be the correct assumption for natural languages or any languages that limited agents like human beings are able to learn. The worldmaking language need not be so restricted, however, for no human being needs to be able to learn it. We are limited and perhaps no language that we can learn can have enough sentences to represent all the possibilities, but what we need is just to find such a way of representing, whether or not any agent can use such a system of representation. The worldmaking language can have, then, infinitary sentences.

Finally, it seems we want the worldmaking language to have the same structure as reality itself and we want the semantic values of the terms of the language to correspond to the elements of reality: the various individuals, properties and relations (Jago (2015, p. 593)). We can, then, take the worldmaking language to be a lagadonian language, a language where every individual is a name for itself and every property and relation is a predicate denoting itself. As Jago argues, taking these terms to be anything but the bits of reality themselves seems to just introduce unnecessary complexity to this picture. Here it is of note that worries about learnability and practicality have already been put to the side, by abandoning the requirement that this be a language that anyone could use. We can, then, take our worldmaking language to be a lagadonian language allowing for infinitary sentences, which is to say, infinitary constructions out of properties, relations and individuals. Taking our worldmaking language to be a lagadonian language also allows for an intuitively plausible explanation of why certain modal claims are true. Let us go back to the English sentence S from before, and let S* be now the translation of S into the lagadonian language, so that S would be true because S* is true, where S* would be something like <<Donald Trump*(the individual), lost the election* (the property of losing an election)> and* (the connective) <Donald Trump* wins the election*>> $\in w$, where w is a maximally consistent set of sentences of the lagadonian language[41]. This means that the truth conditions of the English sentence are given by combinations of the parts of reality that the sentence is, intuitively, about[42], which seems to be the right result, as what seems to make a given sentence true[43] is that certain objects have certain properties or are related in certain ways (if it is right that reality can be exhaustively divided into individuals, properties and relations), which, in the case of possible worlds, will have to be the case in a different situation that decides every question and is in that sense maximal, and which must be consistent (for otherwise it would not be possible). We can then provide a non-circular explanation of the truth-conditions of S in terms of its translation to the lagadonian language.

It seems, then, that appealing to an infinitary lagadonian language deals with both the circularity and cardinality objections to proper linguistic ersatzism. As it was seen, there are, of course, some worries in regards to classifying the lagadonian language as a language: for instance, it seems that no human being can learn it or speak it. Of course, this problem already poses itself when we make the move from a finitary language to one that allows for infinitely long sentences. Even so, they are still languages in the formal sense that they are symbolic systems of representation. In the present discussion of constructivist ersatzism we have tried to present a view that provides enough entities to represent all the possibilities, and which should be suitable to play the role of possible worlds, given a plausible account of their nature. For this it only matters what the entities are and that we have enough of them, just as in the case of the other views so far considered - for instance, we wouldn't need to be be able to grasp in thought or language all of the states of affairs making up the different possible worlds, and it is not for that reason that we have put the view of worlds-as-states-of-affairs to the side. What possibilities there are and what possibilities we are able to grasp and refer to are distinct matters. We can, then, present a view that meets the various objections so far considered: the view of possible worlds as maximally consistent sets of lagadonian language sentences, which can include for instance Russellian tuples, while also containing the usual connectives, quantifiers, variables[44] and operators from modal logic, alongside the individuals, properties and relations.

[41] Here the modal logic S5 is being presupposed, so that each possible world accesses every possible world. This assumption is justified as the modalities under discussion are metaphysical and logical possibility/necessity, for which S5 is widely accepted.

[42] The notion of aboutness will be more carefully considered in the second part of the present dissertation. Here the claim is merely about the pre-theoretical everyday conception of what sentences are about. See Williamson (2007, p. 26) for an explicit appeal to such a notion.

[43] In the second part of the dissertation we will likewise explore further some topics in truthmaker semantics.

[44] Here a difficult puzzle, for which I have no solution at this moment is how to account for variables in a lagadonian language, as given how different the language is, it is not obvious what the variables would correspond to, whereas it is much more clear to consider what the names and predicates of the language would be (the actual individuals, properties and relations).

The problem of aliens: alien objects

It seems, then, that the best version of linguistic ersatzism is one that takes worlds to be sets of sentences in a worldmaking language, a worldmaking language whose atomic sentences can be taken to be Russellian tuples. We now turn to two general objections to constructivist ersatzism, and try to address them with the resources of this latter view: the objection from alien objects and the objection from alien properties, where "alien" just means "non-actual". The objections have the same form: the actual world only has certain objects and properties, and there are only so many of them; but there could have been more objects and properties, or there could have existed different objects and properties (even if less in number) than the ones there actually are; however, the abstract constructions that the linguistic ersatzist can resort to are made up from actual objects and properties; so linguistic ersatzism does not have enough resources to represent all the possibilities, since it seems that some such possibilities contain non-actual objects and/or properties.

Importantly, linguistic ersatzism does not appeal only to the more familiar actual objects and properties in the construction of its worldmaking sentences, but also to the connectives, quantifiers, variables and operators from quantified modal logic. So it is true that linguistic ersatzism represents possibilities via constructions out of actual objects and properties, but the possibilities that it can represent are also made up of the aforementioned resources from quantified modal logic. For instance the possible proposition that there is something can be represented by $\exists x \exists y(x = y)$ which does not contain any actual entities besides the ones we take to be the quantifiers and variables, as well as the relation of identity. The question then becomes whether linguistic ersatzism can represent all possibilities containing non-actual objects and properties with the resources it has left: the resources from quantified modal logic.

Let us start by considering the objection from non-actual objects and by considering the following example (from Fine (2005), p.164). Suppose that a given radioactive material did not emit any particles in the actual world at a given time but that it could have emitted two particles. Let us suppose further that these particles are intrinsic duplicates of one another (as in the earlier example involving electrons) and that they follow different trajectories upon being emitted by the radioactive material. Let us call such particles *alpha* and *beta*. Let *w* be a possible world in which *alpha* follows trajectory *t*1 and *beta* trajectory *t2*. Finally, let *w'* be a different possible world that is exactly like *w*, save for the fact that in *w*' *alpha* follows trajectory *t2* and *beta* trajectory *t1*. Intuitively, these correspond to different fully determinate possibilities. Our theory of modality should, then, treat them as distinct possibilities. However, given that neither *alpha*, nor *beta* actually exist, they cannot figure as names for themselves in any sentences of the lagadonian language. It seems that we would then have to refer to them using other resources from the language. This, however, does not seem possible as the only difference between *w* and *w'* is that *alpha* and *beta* switched places, so that it seems that everything else should come out the same in regards to both worlds. Relevantly, definite descriptions will not help us as it seems that any definite description that will pick up *alpha* in one world will pick *beta* in the other, and vice versa, so that it seems that only appealing to definite descriptions we could not distinguish between situations where *alpha* and *beta* switched positions and everything else remained the same.

One important feature of this example and objection is that it presupposes haeccetism, the view according to which there is for every individual the property of being that same individual, which might also be called its haecceity, so that two worlds which are qualitatively indistinguishable may correspond to distinct possibilities, if two individuals have "swapped" all qualitative properties. So for instance according to haecceitism Obama could have in another possible world all the properties that Sider has in the actual world, and vice-versa, so that the resulting non-actual possible world would be indistinguishable from the actual world but nonetheless correspond to a distinct possibility[45]. Given that in the example above *alpha* and *beta* switch positions but everything else remains the same, then the two possible worlds would correspond to distinct but qualitatively indistinguishable situations. A linguistic ersatzist could simply reject haeccetism and therefore reject the objection on those grounds: the supposedly different possibilities are not really distinct, so no distinct possibilities are being conflated. The problem with such a defense, however, is that it commits a view about the nature of possible worlds with a controversial metaphysical claim which, if possible, it should remain neutral on. Given that anti-haeccetism is a controversial metaphysical position, a defense that instead faces the challenge head on should be preferred, as long as it does not bring with it a greater theoretical cost.

Jago (2014, pp. 152-153) holds that such a position is tenable, if we appeal to Kaplan's (1978) DTHAT operator. This operator takes a definite description, *d*, that, if satisfied, uniquely picks out the object, *o*, that has some properties, *P1, ..., Pn*, and returns *o*, without attributing *P1,..., Pn* to *o*. Intuitively, we can think of the DTHAT operator as switching a definite description from its standard attributive form, to a non-attributive or purely referential form. Jago's idea is then that we can take a definite description to pick out *alpha* in one world and *beta* in the other, for instance the definite description *the particle emitted that follows trajectory t1* (henceforth *T*) and then use the operator DTHAT on it, so that we would get DTHAT(*T*). Since *T* picks out *alpha* in *w*, DTHAT(*T*) will likewise refer to *alpha*, but without attributing to it any particular trajectory or property. DTHAT(*T*) is a rigid designator, so we can use it to refer to *alpha* in all possible worlds and therefore we can use DTHAT(*T*) as a name for *alpha*, occurring in lagadonian sentences about *alpha*. It is not easy to see to what the DTHAT operator would correspond to in a lagadonian language, but perhaps it could be taken to be a function from a definite description (in the lagadonian language) to the unique object satisfying it (when there is such an object)[46].

[45] See Sider (2020, p. 4 and following) for a further development of this suggestive example, as well as for a discussion of haecceitism in how it relates to different concepts of individuals, for instance as bundles of universals or as bare particulars.

[46] Here I would like to thank Franz Berto for the suggestion that some of the terms of the lagadonian language can correspond to functions (for instance the connectives could be taken to be functions from truth-values to truth-values, and the quantifiers to higher-order functions), which prompted in me the thought that likewise the DTHAT operator could be taken to be a function.

A problem for this solution, however, is that since linguistic ersatzists argue from an actualist point of view, they cannot attribute existence to *alpha* or *beta* as, by hypothesis, they are both non-actual particles. But if they do not exist, and if, as assumed, the radioactive material does not actually emit any particles at the relevant moment, nothing seems to uniquely satisfy the definite description *T*. One might grant this point but claim that even if the definite description is not actually satisfied by any object, it would be satisfied had *w* been the actual world, case in which *alpha* would exist and be the object that satisfies *T*, and that this is all that is needed for the definite description to do the work it is supposed to do.

It is not, however, clear, that the same applies to DTHAT(*T*). For while the semantic value of *T* seems to be given by a bundle of properties, the semantic value of DTHAT(*T*) seems to be the object that uniquely satisfies *T*. If the semantic value of DTHAT(*T*) did not contain *alpha*, then it would contain other individuals and/or properties and relations. If it is to be a name for *alpha*, it cannot contain any other individuals, and if it contained any properties and relations, then it seems that its semantic value would just be the same as *T* and that it wouldn't refer to *alpha* in all possible worlds, as per hypothesis *alpha* and some other particles, like *beta*, are intrinsic duplicates of one another, or at least they are so in respect to the actual properties[47], so that there is no intrinsic property that one of them has in all possible worlds that all the other particles lack, and likewise there is no extrinsic property that any object has in all possible worlds (such as following a given trajectory) that no other object has. Finally, if the semantic value of DTHAT(*T*) included more than *alpha*, then it seems that at least to some extent DTHAT(*T*) would, like *T*, be attributive. DTHAT(*T*)'s semantic value, then, seems to just be *alpha*. So that since *alpha* does not exist, it seems that DTHAT(*T*) does not have a semantic value and we are not able to, therefore, use it as a name for anything in the lagadonian language. It does not seem, then, that using Kaplan's artificial operator will help the linguistic ersatzist with the problem of alien objects. Intuitively, it seems that just like with the English demonstrative *that*, in the case of DTHAT we need to in some way be given the object uniquely picked out by the definite description in order to then use DTHAT to single it out and refer rigidly to it. The claim is, basically, that the present proposal is akin to pointing to the void and saying "that", case in which, no object being given, the speaker has failed to refer.

Even if Jago's solution does not work, I believe that the linguistic ersatzist may still insist on using definite descriptions as surrogates for non-actual but possible objects. This move shares some features of Plantinga's strategy, as non-actual objects still have actual surrogates given in terms of certain properties. However, while Plantinga helped himself to necessarily existing properties of being a given object, here the claim is instead that the surrogates for non-actual objects are given instead by certain combinations of actual properties, these last properties being instantiated by some actual objects, which had the world been relevantly different, would be

[47] Here I am assuming that either there are no such properties as *being-alpha*, or that such properties, or haecceities, are not necessary existents. If there were, the problem would be defused from the start as the property *being-alpha* would be an actual property and we could, therefore, distinguish *alpha* and *beta* with the help of their respective haecceities. This is Plantinga's (1974) solution for the problem of non-actual objects and properties, which has been faced with (what I take to be) decisive objections (see for instance Stalnaker (2012) and Menzel (2014) for objections to this perspective).

uniquely instantiated by the non-actual object they serve as a surrogate for. In the earlier strategy, as I'll now argue, what was doing the work in terms of picking out the same object across all possible worlds did not have anything to do with the non-attributive nature of DTHAT per se, but rather with the fact that there is an implicit world-relativization. First, consider the fact that one definite description can be uniquely satisfied by different objects at different possible worlds, and that this is the case for *T*, which is satisfied by *alpha* in *w* and by *beta* in *w'*. Second, and crucially, consider that it is only relative to a possible world where a definite description is satisfied that DTHAT is defined. In fact, DTHAT won't pick any objects if its sole input is a given definite description, for what object will satisfy the given definite description depends on what world we are to evaluate the definite description relative to. It seems, then, that we should add an argument for a possible world, so that DTHAT would be a function from a definite description and world to an object. Third, consider that it is by fixing on a possible world, that is taken as an argument in the function, that we can then retrieve an object, to which DTHAT(*T*, *w*) would rigidly refer.

Since following the suggestion of using DTHAT led us to a problem for the linguistic ersatzist, we might, then, want to explore a different way of fixing on a possible world and using a definite description to represent the distinct possibilities at hand. One way to do so is to take the possible world relative to which the DTHAT operator would be defined not just as a separate argument, relative to which the definite description is defined but rather as part of the definite description itself. In the case of the example so far considered, instead of using *the particle that follows trajectory t1*, we could use *the particle that follows trajectory t1 in w*[48]. In both cases *alpha* is the object picked out by the definite description, but in the second case it is so in every possible world, for it is necessarily true that *alpha* is the particle that follows trajectory t1 *in w*. This latter definite description, then, would refer rigidly to *alpha* and would, therefore, allow us to distinguish between the worlds at hand, for even in *w' alpha* is the unique object satisfying the description: *the particle that follows trajectory t1 in w* (henceforth *TW*).

Here an objection might be raised in regards to the well-foundedness of the relation of set membership. It seems that at *w*, worldmaking sentences containing *TW* should be true, for instance the following: P(*TW*) [to be read as *the particle that follows trajectory t1 in w is a particle*]. Truth in a world, on this view, however, corresponds to world membership, so that P(*TW*) $\in$ *w*. However, *w* itself is part of *TW*, so it seems that in order to use the aforementioned world-relative definite description to single out *alpha* in all possible worlds, we would then be committed to *w* $\in$ *w*, which violates the axiom of regularity, that has as a consequence that no set is a member of itself. The current proposal would, then, violate one of the axioms of standard set theory. I believe that such an objection would be misguided, as it does not follow from P(*TW*) $\in$ *w* that *w* $\in$ *w*, for even if *w* figures in *TW*, the relation between *w* and *TW* is not one of set-membership, and it is not in general the case that if s $\in$ s' and s' $\in$ s'' that s $\in$ s'' for sets s' and s'' and any object s (set or not): Socrates belongs to the set of people, the set of people belongs to the set of numerous sets, yet Socrates himself does not belong to this latter set (Socrates is not

[48] Given the nature of the worldmaking language here accepted, definite descriptions of this form will be terms of a sentence which is a member of a world, which is itself a constituent of the definite description. This leads to some circularity worries, which we move on to discuss.

a numerous set). One might argue, as Lewis does against impossible worlds, not from the relation of set-membership, but from the relation of parthood, where transitivity more plausibly holds (if *p1* is part of *p2*, which is itself part of *p3*, *p1* is part of *p3*). Since *w* is a part of *TW*, *TW* a part of P(*TW*) and P(*TW*) a part of *w*, then we have that *w* is a part of *w*. By itself, this result is not problematic as we can accept that everything is a part of itself, albeit not a *proper* part of itself. The problem, rather, is that it seems that *w* would be part of a proper part of *w*, which goes against the intuitive notion that the proper parts of a whole should not contain that whole as a part, for wholes include all of their proper parts, but are not included in any of them. The relation of proper parthood is antisymmetric.

The objection stated in terms of the relation of parthood seems, then, to be much more promising than in terms of the set-membership relation. It is not, however, clear that it is successful either. First, the objection requires both that the relation of parthood is transitive and that sets have their members as parts, but the conjunction of these theses is untenable. Suppose that *C* is the set of all cats and that every cat has whiskers. Since a given cat's whiskers are parts of that same cat, and the cat part of *C*, by the transitivity of the relation of parthood we would have that the whiskers are part of *C*. Further, since the parts of *C* are its members, we would then have that the whiskers would be a member of *C*. *C*, however, is the set of all cats, so that all and only its members are cats: only the cat to which some whiskers belong to is a member of the set, not the whiskers (by) themselves. We cannot, then, hold at the same time that the relation of parthood is transitive, and that the members of a set are its parts.

Second, it is by no means clear, given recent developments in mereology, that the relation of proper parthood ought to be, or at least ought to be in all cases[49], antisymmetric. Cotnoir and Bacon (2012) present a number of cases where this condition seems to fail, among which is the following. Suppose that "The universe exists" expresses a proposition, and that "universe" refers to all there is, not just concrete objects. On the Russellian view of what propositions are, the universe itself would be a constituent of the proposition, but the universe contains all there is, and so it also contains the proposition expressed by "The universe exists". We have, then, that the proposition is part of the universe, which itself is part of the proposition, and yet the proposition and the universe are distinct, so antisymmetry would fail. Of course, one could take this example to be a further argument against Russellian propositions, not against the antisymmetry of the relation of parthood: as so often occurs, a philosopher's *modus ponens* is another's *modus tollens*. Even so, just as we have accepted the claim that a view about the nature of the possible worlds should not settle the matter in regards to whether haecceitism or anti-haecceitism is right but should remain neutral on the matter, perhaps the same could be said about the nature of parthood precluding certain positions in the debate over what the nature of propositions is.

---

[49] Here this caveat is presented to allow for the view that there are various relations of parthood, each of them with its own properties and structure. See for instance Fine (2010) for a defense of this view. The following are examples of relations thought to correspond to different notions of parthood: material constitution, functional constitution and abstract composition. See for example Thomson (1998, p. 155) for a defense of the position that *a* constitutes *b* if and only if *a* and *b* are parts of each other but not identical. This is, then, a view according to which the relation of proper parthood is not antisymmetric.

Third, it seems that if the argument was sound, then it would equally undermine all ersatzist approaches, insofar as in all of them the possible worlds are taken to be abstract and all to actually exist. Some examples of ersatzist positions are the following: possible worlds are maximal properties; maximally consistent states of affairs; maximally consistent sets of propositions; maximal propositions and maximally consistent sets of worldmaking sentences. It also seems to be the case that if a given proposition, *p*, is true in a world, *w*, then the proposition *p is true in w* is true in all possible worlds (since we're assuming S5), including the ones that *p* is true in. So it'll be the case in *w* that *p is true in w*. More generally, it seems that some necessary propositions are about worlds, for instance: *w is a possible world*, where *w* is indeed a possible world. If worlds are not concrete objects as in Lewis (1986), then it seems we will run, on all of these accounts, into the earlier result that (again, accepting the transitivity of parthood and that the members of a set are parts of it) a world is part of a proper part of itself. To illustrate, let us consider the case of the view of possible worlds as maximal properties. On this view, it seems we can think of world-properties as being complex properties formed out of other properties. Just like the property *being F and being G* is a property possessed by anything that is both *F* and *G*, we can think of possible worlds as being properties made out of more basic properties in the same way. Intuitively, the properties contained in this way in the maximal property that corresponds to a given possible world will be parts of it. This world-property, *P*, will be maximal when, just like in the case of worlds as maximal states of affairs, for a given pair of incompatible properties, *P1* and *P2*, either *P1* or *P2* are included in one or more conjuncts of *P*. Now, we have two options, depending on whether we take propositions to be sets of worlds or structured entities. If the proposition in question, *w is a possible world*, is a set of worlds, then *w* will be part of the proposition (since we are assuming that the members of sets are parts of sets), and the proposition in question will, in turn, be part of a property that is part of *w*, so that, by the transitivity of the relation of parthood, the proposition will be part of *w* as well. If, on the other hand, the proposition is taken to be a structured entity, then plausibly it will also contain *w*, as in the Russellian view, so that again *w* will be part of the proposition, and the proposition in question part of *w* (since it is a part of a property that is part of *w*).

Mutatis mutandis this argument structure also applies to the view that takes possible worlds to be maximal states of affairs or maximal propositions. As for the views that take worlds to be maximally consistent sets of propositions or states of affairs, likewise the argument presented earlier against proper linguistic ersatzism could, with suitable changes, be extended to them. Of course, that the argument is so general does not show that it is misguided, just like Lewis's argument against magical ersatzism considered earlier. However, as in that case, it seems that it puts higher stakes on it, for we could take the argument to show instead that there are a variety of new cases that give force to the position that either our mereology should not, in these cases, be well-founded, or that the members of a set are not parts of that set[50].

[50] An important asymmetry with the case of Lewis's argument is that if Lewis's argument would be turned on its head, then instead of showing that we do not actually understand certain relations (the selection relation and of set membership, as Lewis claims, and, if the earlier arguments were right, also relations of ontological dependence between abstract simples and concrete objects and of *being-a-mode-of-presentation-of*, between simple Fregean senses and referents of singular terms), it would prove that there are relations which are neither internal, external or

It seems, then, that the main worry about having *w* be a constituent of one of the worldmaking sentences that constitute *w* has been dealt with and we can proceed in using the world-relative definite descriptions to stand for the non-actual objects, enabling us to make the distinctions between possible worlds that the haecceitist maintains there are. An important point to take notice of is that the view so far presented is not *committed* to an haecceitist view: it merely provides the framework on which haecceitist distinctions can be made, if they are wanted. We now move on to consider the objection from alien properties, providing a similar solution to it.

The problem of aliens: alien properties

Intuitively, just as there could have been more or different objects than those there actually are, there could have been more or different fundamental properties than those there actually are. Here the focus on new *fundamental* properties stems from the fact that, had these properties not been fundamental, then perhaps they could be described in terms of actual properties, case in which the ersatzist could use the latter to describe the possibilities involving the "new" properties. If the new properties were fundamental, then by hypothesis they could not be expressed in terms of other properties, for it would be *these* latter properties that would then be fundamental and not the properties that we started with. If one rejects a distinction between fundamental properties and non-fundamental properties, then instead of fundamental properties we could run the example instead with properties that we have assumed are not equivalent to any combination of actual properties. The objection is, then, again that constructivist ersatzists cannot account for all the possibilities, as they have no way of representing the relevant non-actual properties.

Two important features of this objection are the following. First, it presupposes that there are no uninstantiated properties. Second, it seems that no example can be given by the objector of an alien but fundamental property. The first point follows immediately from the example: assuming that there could have been more or different properties than there actually are seems to assume that properties are not necessary existents, and the properties that exist at any given world would then be the properties that are instantiated in that world. The second point in turn corresponds to the thought that to be able to present an instance of a fundamental alien property one would need to be acquainted with any, but if uninstantiated properties do not exist and properties only exist in the worlds in which they are instantiated, then alien fundamental

---

extrinsic. But these three options seem to cover the entire space of possibilities for what kinds of relations there are, in terms of their division according to what kinds of properties they hold in virtue of (intrinsic, extrinsic, or a combination of the two, in the case that an extrinsic relation can be broken down to the combination of an internal and an external relation, as I argued that the relation *has-a-length-in-metres-of* is). In the present case, however, turning the argument on its head would prove either that in at least some instances the relation of parthood ought not to be treated as both transitive and antisymmetric, or that the members of a set are not parts of it, instead of proving that all ersatzist approaches considered are misguided because worlds cannot be parts of proper parts of themselves. Contrary to the case of the division of relations into internal, external and extrinsic, it seems that there are alternatives to both the thesis that sets have their members as parts, and that parthood is both transitive and antisymmetric, which have been independently motivated (see the first two points in reply to the objection), which do suggest that the argument is not sound.

properties do not exist in the actual world nor can they be aptly described in terms of properties existing in the actual world. Nothing, however, will turn on this last point that no example can be given of a fundamental alien property.

Like in the case of alien objects, one way for the ersatzist to reply would be to just deny that there are no uninstantiated properties, or that properties are contingent beings. But just as the commitment to anti-haecceitism was thought to be unwelcome, given that it is a controversial metaphysical claim that our view of the nature of possible worlds should, if possible, avoid, likewise the view that there are uninstantiated properties is a controversial metaphysical claim that, it is argued, we would do best to avoid (see Armstrong (1978, chapter 7) for an influential rejection of the existence of uninstantiated properties). It seems, then, that in order to remain neutral on the debate of whether there are uninstantiated properties, the ersatzist should provide us with tools to represent the relevant non-actual properties.

One immediate thought is to use a property's nomic role ro represent it. This strategy mirrors the strategy of representing non-actual objects by the properties they have in another possible world via a definite description that is not satisfied, but would be satisfied had the possible world in question been actual. A given property's nomic roles can be represented by all the sentences which it makes true. In the case of non-actual properties, given that the property does not exist and therefore cannot serve as a name for itself, we can use a Ramsey sentence to represent its nomic role: $\exists X \varphi X$, where X is a second-order variable and φ is a formula corresponding to the conjunction of formulae that are true of that property in the relevant possible world. We can then say that *P* has nomic role *R* if and only if, when eliminating the second-order existential quantifier and replacing every instance of X for *P* we get the true sentence φ (*P*). Finally, we can represent the non-actual property *P* by the definite description: *the property with nomic role R.*

Here an immediate objection mirrors the objection to the corresponding solution for the problem of alien objects. Just as objects might have different properties in different possible worlds, properties might also have different nomic roles in different possible worlds, so the definite description *the property with nomic role R* will pick out different properties in different possible worlds. We wouldn't, then, be able to distinguish between possibilities where the only difference is that two properties have exchanged nomic roles and our conception of what possible worlds are would then preclude quidditism (that is, the view according to which two properties can swap nomic roles, leading like in the case of haecceitism for individuals, to two indistinguishable possibilities) in favour of nomic essentialism (the view according to which what properties are is entirely given by their nomic roles, so that if two properties swapped roles, the resulting situation would just be the starting situation and there would not be *two*, but only *one* possibility), in the same way that earlier the solution for alien objects favoured anti-haecceitism over haecceitism.

Again mirroring the solution for the case of alien objects, we can relativize the definite description to the world, so that we would get: *the property with nomic role R in w*. This definite description will then refer to the same property in all possible worlds. Bar any unforeseen arguments for appealing to definite descriptions mentioning nomic roles[51], and having dealt with the objection for using "*in w*" as a part of the definite description, then it seems that definite descriptions of this form could be used to denote fundamental properties, so that in this framework mere quidditist distinctions between worlds can be accounted for.

[51] One objection that could be raised at this point is that this view implicitly accepts that the totality of nomic roles is already given, and so that there could not have been more nomic roles than there actually are. This is indeed a commitment of the present view, however it is not as substantive a commitment as it might at first seem to be. Given that the nomic roles of non-actual properties can be given by descriptions containing second-order variables, the nomic roles effectively correspond to ways for properties to be instantiated by objects and how they combine with other properties. But even if there could have been more properties than there actually are, there seems to be no reason to suppose that the possible combinations of properties cannot be given, even in the absence of the properties themselves.

Pending questions and objections

Having presented various forms of linguistic ersatzism, and having argued against some of the major objections to some particular views, as well as to the general approach, I believe that a favourable case has been presented in favour of proper linguistic ersatzism. In this section I would like to address some loose points and remaining issues that have been lying under the surface in the preceding discussion.

Singletons, parthood and Finean mereology

It was concluded earlier that the best way for a constructivist ersatzist to construct worlds would be out of concrete objects, for otherwise it would seem that Lewis's objection against magical ersatzism would also apply to the various constructivist solutions. It was also seen later that taking the relation of parthood to be transitive and members of sets to be their parts we would reach the wrong conclusion about what the members of a given set are. But here there is a dilemma for the view so far exposed. If sets do not have their members as parts, and if possible worlds are sets of worldmaking sentences, then it seems that, contrary to what was stated, we cannot maintain a view of possible worlds where its ultimate constituents are concrete, and we will thereby run into the same argument Lewis has presented against magical ersatzism, as sets seem to be abstract objects, at least in the sense that the ersatzist approach so far considered takes worlds to be abstract constructions out of actual objects and properties. In fact, Lewis harshly criticizes the notion of a singleton for the reason that it is not clear why it is one singleton that is the singleton of a given object and not any of the other singletons, which, in his view, are all mereological simples. On the other hand if we do accept that the members of a set are its parts, then it seems we would have to reject the transitivity of parthood, which again seems to be a controversial claim that our theory of what possible worlds are should not be committed to.

Perhaps we can avoid the dilemma and we can have that the members of a set are parts of the set while also holding on to transitivity. One first attempt to do so would be to reject that only the members of a set are parts of it, and claim that even if a given cat's whiskers are not members of the set of cats, they are still part of it, given that they are part of a cat that is part of the set. It would be hard, however, to see why one would accept this claim if one was not already committed to sets having their members as parts and to the transitivity of the relation of parthood. This reply would, therefore, presuppose the very matter that is at issue. But perhaps we could say in its favour that with some modifications, it could be used to make sense of the set membership relation, which, as seen, may not be properly understood at all, for it is unclear what the nature of the relation between an individual and its singleton is. The modification would just be to introduce the notion of a "first part". We could say that for a given composite object C (either individual or class) and object x, such that x is a proper part of C, x will be a first part of C if and only if for all y such that y is a proper part of C, x is not a proper part of y (that is, x will be a proper part of C, but it won't be a proper part of any proper part of C). We could then identify set membership with the first-part relation, avoiding the objection presented earlier that if this were so, then the set of all cats would have cat whiskers as some of its members, which it does not.

Such an approach would, however, not work, as Fine(2010, p. 565) shows with the following example. Consider the set composed of x and x's singleton: {x, {x}}. Let us call it X. If the current understanding of membership in terms of first parthood was correct, however, then x could not be a member of X, for x is a proper part of {x}, and {x} a proper part of {x, {x}}, so that x is a proper part of a proper part of X, which, by definition, precludes x from being a first part of X.

Is there a way of remedying the notion of a first part so that we could still define membership in terms of the relation of parthood? Here's an attempt to do so. Intuitively, what makes the aforementioned definition lead to the wrong result that x is not a first part of X is that we imposed that x *must not* be a proper part of any proper part of X. The difficulty, then, becomes how to define first parts in a way that allows us to again distinguish the members of a set from its other parts, while allowing first parts to, in some cases, *also* be proper parts of proper parts. We might do this by resorting to the notion of a chain of proper parthood. Let P be a proper parthood chain between x and C when there are some y1, y2, …, yn such that x is a proper part of y1, y1 a proper part of y2, … and yn a proper part of C, so that we can say that the ys link C and x together. We can then define a first part as follows: x is a first part of C if and only if x is a proper part of C and there is a proper parthood chain P ending at C and in which x figures such that for all y distinct from C that are in P (that is, for any element of the chain that is not at the end of the chain), x is not a proper part of y[52]. According to this definition, to be a first part of a given object C is just to be the penultimate element of a chain ending in C. Alternatively and informally, we can define a first part as a proper part of an object such that it may be the case that it is a proper part of that object without it being a proper part of any proper part of that same object, that is, it may be that no proper part of the object "gets in the middle" of the first part and the object considered, relative to the relation of proper parthood[53]. Similar definitions could then be given by iteration for putative second parts, third parts and so on.

A problem for this strategy, however, is that it does not seem to make any progress in terms of how to understand the relation holding between a singleton and its member, for it would

---

[52] Here we could have just as easily defined a chain in terms of parthood instead of proper parthood. Taking that route, however, would mean that chains could contain certain objects multiple times over. Constructing chains in terms of proper parthood helps trim down the chains.

[53] Here "may" is not to be read modally, but rather simply as elucidating that there are certain cases corresponding to such chains, which are not the only chains of proper parthood, so that it is not always the case that a first part is a proper part of a given object without being a proper part of a proper part of that same object at the same time. This means that in some cases, as the one seen earlier of {x, {x}} and x, there is more than one chain P which links C and x together: what is needed is simply that in one of these chains x is a proper part of C that is not a proper part of any y that is a proper part of C. If we give up the anti-symmetry of the relation of proper parthood, then it also may be the case that the chain P contains loops of proper parthood. But even so, x will still be considered to be a first part of C as even if cyclical there is a chain in which x is a proper part of C without being a proper part of any proper part of C. In the case of a very close loop where C is a proper part of x and x a proper part of C, where both x and C meet the condition just presented, both x and C are first parts of one another.

be the case that the singleton of Socrates, {Socrates}, would contain, on this proposal, as its only parts Socrates and all of his parts. Here, Socrates himself would be the singleton's only first part. A given object only having one first part seems, however, to go against the intuitive principle that whenever x has a proper part, y, there is a z that does not overlap y and which is a part of x, that is, the principle of weak supplementation, also often referred to as the principle of the remainder. In the case of the singleton, however, all parts of {Socrates} which are not themselves Socrates are nonetheless parts of Socrates, and therefore overlap Socrates. It seems then that in this case we have merely traded in one mystery for another: the mystery of what the membership relation is for the mystery of how it is that singletons only have one first part.

The suggestion so far explored is similar to Fine's (2010) reply to the present concerns. He suggests that for x to be a part of a set S is given not in terms of the membership relation itself, but rather in terms of the ancestral of the membership relation, that is, a given x will be a part of S if and only if x is a member of S, or a member of a member of S, or a member of a member of a member of S, and so on, being the members of a set just its "direct parts", which correspond to what we have called the "first parts" of S. There are, however, important differences between Fine's reply and the reply sketched above. For instance we claimed that all parts of Socrates are parts of his singleton, but that's not so in Fine's view, for nothing, including Socrates's parts, is a member of Socrates. So, according to Fine, only Socrates himself, but none of his parts, is a part of his singleton. One might want to accept this hypothesis instead if one takes it to be more intuitively plausible and to avoid the ad-hoc worry in regards to the first parts hypothesis, that is, the worry that there are no independent reasons to maintain that the parts of a member of a set are themselves parts of the set. However, by accepting such a move, it would seem that we would have to give up on the transitivity of the relation of part, for a part of a part of a set won't in general also be a part of the set.

A different aspect of Fine's view is that the way of composition characteristic of sets and the way of composition by which material objects are formed from their constituents are inherently different. In the case of the former, the wholes that get formed are highly structured in an hierarchical level - for instance we can think of an individual as occupying the level 0, the singleton of that individual as occupying the level 1, and so forth for each application of the set-builder operator. In the case of the latter, every individual formed out of the fusion of other individuals is, nonetheless, just an individual as well, so that this form of composition is "flat", in the sense that forming new individuals by mereological fusion does not create new levels in a hierarchy of structural complexity[54]. To make this explicit, consider the following examples of flat and hierarchical composition. For Lewis classes that are not singletons are simply fusions of singletons, including classes such as the set {x, {x}}, which is the mereological sum of the classes {x} and {{x}}. Here, all classes are on the same level, as they don't add any "unmereological way of composition"[55]. Fine, on the other hand, prefers a view of sets that he labels "operational", as sets in his view are primarily understood in terms of the successive

[54] Fine (2010, p. 566).

[55] See Lewis (1991, p.40 and pp.55-56) for an explicit rejection of "unmereological composition".

application of the set-builder operator. So where *a* is any individual, SB(*a*) will be the set formed by one application of the set-builder operator (*a*'s singleton), SB(SB(*a*)) will be the set formed by two applications of the operator (*a*'s singleton's singleton), and so on. The empty set and all pure sets then correspond to the limit cases where the set-builder operator is defined on zero objects, for instance the sets SB() and SB(SB(),SB(SB())), which are respectively {} and {{},{{}}}. On this approach, the set {x, {x}} has a higher level of structure than, say, the set {x, y}, for while the latter is just SB(x, y), so that it corresponds to only one application of SB(), to get to {x,{x}}, two applications are needed, as it corresponds to SB(x, SB(x)). In Lewis's view, since {x} and {{x}} are both mereological atoms - like all singletons -, the set X corresponds to the fusion of two singletons just as {x, y} does. So in Fine's but not in Lewis's view, {x,{x}} has a higher level of mereological complexity than {x, y}.

Fine's approach is surely interesting, but it's not clear how it helps with the problem of how to make sense of the relation between an individual and its singleton. If we're talking about the mere writing and manipulation of symbols, then it's clear how the set-builder operator behaves: it takes a given string of symbols as an input and outputs that string now with the symbols "{" and "}" flanking it. But the current discussion is not about mere symbols and their manipulation, but rather about what sets and their parts are, that is, about their metaphysics. Fine's operational strategy seems to tell us how to get sets from their members and what the parts of the set are (the objects that serve as input for the set-builder operator[56]). But in the case of the singleton, the only part of a singleton will then be its only member, for the set-builder operator will only be applied to one object. This, however, commits Fine to the rejection of the principle of weak supplementation, and it may, then, seem that one has simply traded one mystery for another, that is, the mystery of how it is that singletons and their members are related for the mystery of how there can be a whole composed of only one proper part.

Mereologies where the principle of weak supplementation does not hold have been proposed, and some examples to illustrate cases where one might reject this commitment of classical extensional mereology have been presented. Such examples include Yablo's (2015) case of the difference between a sphere and its interior, where despite the sphere containing its interior, there is nothing to make up the difference between the interior and the sphere, as well as the much discussed example of the clay statue - case in which, according to some, the clay is a proper part of the statue, as it is not identical to the statue, but there is no further thing to make the difference between the clay and the statue[57]. We might, then, have reasons to think that the singleton of a given individual just has that individual as a part and nothing else, rejecting weak supplementation. We would then be in a situation where what first was thought to be a very

[56] Fine (2010, pp. 567-568).

[57] Here one might be inclined to state that what makes the difference between one and the other is simply the form or shape of the statue. The form of the statue, however, is not a part of the statue in the sense here considered, where there are no levels of structural complexity, but rather where the fusions are all "flat". Plausibly the form of the statue is not an extra individual that joins with the clay to form the statue.

intuitive principle, even analytic by some[58], later turns out to not be acceptable, on the basis of a more careful analysis. I believe that the preceding considerations may be on the right track, and that we might have good reason to reject the principle of weak supplementation. Yet, if it would be possible to maintain that sets have parts but that singletons have more proper parts than their members (or their members and all their parts), the constructivist ersatzist position so far developed would be more secure in its claim that worlds should ultimately be composed of concrete objects. This option is explored in the paragraphs that follow.

One such way of dealing with the mystery of what the parts of singletons are is developed in Caplan et. al (2010). In it, a non-classical mereology is also adopted, as they take the proper parts of sets to be their members, plus the empty set, which leads to the conclusion that, for instance, {{}} and {} have the same proper parts but are nonetheless distinct sets, contrary to the principle of classical extensional mereology according to which two wholes that have exactly the same proper parts are one and the same - that is, the principle of strong supplementation. Interestingly, however, Caplan et. al (2010) identify the empty set with the property *having some attribute or another*, extending Fine's (1999) view according to which some wholes, including some material wholes, are defined not in terms of simple mereological sums, but are rather structured, being unified by certain objects bearing certain relations to one another or by an object having a given property. The two examples that Fine focuses on are ham sandwiches and cars. Even if we accept with Lewis a principle of unrestricted composition, it does not seem that it is sufficient for an ham sandwich to exist that the ham and the bread it is made of exist at any given time - it seems that the constituents of the sandwich need to be arranged "ham-sandwich-wise". Similarly, some tires, steering wheel, glass, engine and so on do not, dispersed, a car make, but only when they are arranged in a specific way where each of the parts can play a particular functional role in the system. Fine (1999) considers these to be different kinds of wholes in the sense that the slices of bread and the ham are timeless parts of the ham sandwich, but we might remove or change some parts of a car without it thereby ceasing to exist. Yet, they are still wholes where their structure and the relation holding between the parts are relevant, with the latter being a further part of the whole. Here, therefore, I won't consider what Fine has to say about the difference between these kinds of wholes, but rather will focus on the first kind, what he calls "rigid embodiments".

In Caplan et al. (2010) sets are considered to be wholes in just the same way that the ham sandwich is a whole. Further, they take seriously Armstrong's (1991) thesis that for a given objects to form a set, they must be a unit, that is, they must in some way be "a one". That is the reason why they then go on to state that the property corresponding to set-formation is *instantiating some attribute or other*, which is close to Armstrong's state of affairs *having one unit-making property or another*[59]. Following Fine's (1999) notation, we might label the whole

[58] See Simons (1987, p.116) for the claim that weak supplementation is "indeed analytic - constitutive of the meaning of proper part", and Cotnoir (2018) for why it may be that some philosophers take it to be analytic, while others don't even consider it to be true.

[59] Armstrong (1991) agrees with Lewis that classes can be formed from the singletons by mereological sum, so that this general claim, for Armstrong, is equivalent to saying that every member of a set must be a one.

formed by certain individuals *a*, *b* and *c* standing in a relation *R* as *a,b,c/R*, where this is an *object*, not a fact or a state of affairs - it is the ham sandwich rather than the fact of the ham and the slices of bread being in the betweenness relation[60]. The set {a,b,c} would then just be the whole where R is the property *instantiating some attribute or other*. As stated earlier, Caplan et al. (2010, p. 522) go on to identify this property with the empty set, which might be taken to be a counterintuitive move on their part, but if we take sets to just be composed of a material part and a formal part, then it seems that in the limiting case where there is no material part, that is, in the limiting case where a set does not have any members, it should just be identified with */instantiating some attribute or other* which just corresponds to the property by itself. Similarly, we can then apply the property again to get */R/R*, that is, the singleton of the empty set. But no matter what we take R to be, /R and /R/R have the same parts which are just R, so that strong supplementation fails. It seems, then, that even if we want to appeal to Fine's view of wholes that have formal and material parts, we are still left with a non-extensional mereology, where different wholes can have exactly the same proper parts.

Should we, however, accept an extensional mereology? Strong supplementation is perhaps the most controversial feature of classical mereology, and several counterexamples to it have been presented[61]. Yet, whether such putative counterexamples show that strong supplementation is to be rejected is one of the most lively issues in contemporary mereology. Rejecting it would, of course, be a less controversial move for those who want to maintain, like we do, that sets have their members as parts rather than rejecting weak supplementation, which, as the name suggests, is a weaker principle of classical mereology[62]. Still, if we keep upholding a general strategy of avoiding controversial metaphysical claims, it would be good if an answer to the problem of what the parts of sets are that does not deviate from classical mereology was available.

As it turns out, I believe such a move is available, and that both Lewis on one hand and Fine and Caplan et al. on the other got part of the story right about set membership (here I won't suppose that set membership can be defined in terms of mereology and the singleton relation taken as primitive, focusing rather on the relation of set membership itself, which is often recognized as a primitive notion of set theory, instead of the relation between given objects and their singletons). While I take it that singletons are not abstract simples, or mysterious entities, but rather composite objects that have their sole members as one of their proper parts, and that

[60] *Betweenness* is proposed by Fine as the relevant relation for the case of the ham sandwich. This is, as he himself explicitly says, a simplifying assumption. Caplan et al. (2010) explore some difficulties with the proposal. Here nothing I have to claim turns on the correctness of this assumption, so the underlying complicating factors will be ignored.

[61] See Varzi (2016).

[62] This is not to say that one cannot reject weak supplementation while accepting strong supplementation, which would amount to accepting that sometimes objects only have a first part and its proper parts as proper parts while rejecting that there are ever two wholes such that they have all the same proper parts. This is a coherent view, for even if x only has a first part, y, so that y and all its proper parts are proper parts of x, it is still not the case that all proper parts of x are proper parts of y, for there is one proper part of x, y, that is not a proper part of y.

the singleton relation can be made precise, I take it that Lewis comes to an important conclusion in regards to the relation between singletons and their members and what singletons are that more accurately applies in the case of set-membership: that it is not necessary to specify exactly *what* they are both if one is doing mathematics (case in which the notion of a singleton can be taken as a primitive and its metaphysics ignored, for we know enough about its structural constraints) and if one is doing metaphysics (case in which one can opt for a structuralist position according to which there isn't *the* singleton function, but rather various equally good candidates, which meet constraints on their structure and size, so that set theory amounts to the claim that there is at least one such candidate). The basic idea that I want to explore here is that for different sets, the parts that are not members of them will also differ, so that not all sets have the same proper parts that are not members, and thereby avoid the worry that for instance {Socrates} and {{Socrates}} have the same proper parts, which would require that we abandon the principle of strong supplementation. In what follows I wish to make this view more precise.

Sets as fusions of Fine fusions

In order to get a better grasp on what the view I favour amounts to, let us consider more closely the view of Caplan et al. (2010). Following Fine's (1999) remarks about the aforementioned new type of fusion, they take the *candidate* formal Fine parts[63] for the role of being the formal part of the sets to be properties instantiated by all the members of a set. In the case of the set of all women, these might include, among many others: *being a woman*; *being a human*; and *instantiating some attribute or other*. In fact, no matter how dispar the objects we might pick to be the material parts of a Fine fusion, Caplan et al. take it to be the case that there is an attribute that they all have, and so that they also have the property of having some attribute or another. For this reason, they take *instantiating some attribute or other* to be *the* best candidate to be the formal part of every set, and, as stated, this property to be identical to the null set. Besides the fact that there being *one thing* - be it a property, a state of affairs or anything else - that corresponds to the null set leads to the violation of classical mereology, this view is also committed to what might be considered to be a further controversial claim: that there are abundant properties. It seems that in order for there to be enough Fine fusions to correspond to all the sets that set theory says there are, and indeed in order to hold that the predicate "instantiating some attribute or other" corresponds to a property, we must accept the existence of abundant properties. If one does not favour modal realism in the style of Lewis, then one can't define the sparse properties as universals and further the abundant properties as classes of possibilia, which might complicate matters for the view. In fact, even if they accepted modal realism, the theory would then become circular, for sets are being identified with Fine fusions containing an abundant property, and abundant properties are in turn identified as sets of

[63] Here I use "Fine part" to mean, following Caplan et al's use, the material and formal parts of the fusions mentioned earlier (which we might call "Fine fusions") where a new object of the form *a,b,c/R* is formed out of the material parts *a, b, c* and the formal part *R*.

possibilia. Other ways of accepting abundant properties are available, and the issue would require further development, but like we've done so far, we'll ignore such difficulties for now[64].

In regards to the problem of the identity of the formal part of sets, the view I would like to defend rejects that all sets share one and the same formal part, and even further that for any given set there is *the* formal part of that set. Rather than selecting a privileged candidate to be the formal part of a set, we can consider all candidates, contra Caplan et al., to be "born the same way", and take them all to be legitimate. In fact, since no privileged property of the specified objects is going to get picked out, it becomes improper to call them candidates, for there is no position they're "applying" to. There are several ways one might go from here, one of them being to take a given set to be the equivalence class of Fine fusions that share all their material parts. Going back to the example of the set of all women, this set would not be equivalent to *Carla, Maria,.../being a woman, Carla, Maria … /being a human* or even *Carla, Maria .../having one attribute or another*. Rather, the set corresponds to all such Fine fusions, for they contain the same material parts.

The hypothesis that sets are equivalence classes of Fine fusions automatically helps explain some intuitions about how to think of sets in this framework. First, we can account for why the problem of the parts of singletons, and the parts of sets more generally, seemed so intractable - the problem was so prevalent because we kept asking what we needed to add to an individual, say, in order to mereologically form its singleton, whereas, if the current proposal is right, there is no specific thing that, added to an individual, makes up for the difference between an individual and its singleton, as sets are *not* fusions of their members and a formal part to be discovered, but rather are equivalence classes of such fusions. Second, if we accept this view, we can also explain in an intuitive way the difference between the ham sandwich and the set that has as members the ham and the two slices of bread out of which the ham sandwich is made, namely by stating that while for the former to exist the ingredients must be related in a specified way, there are no such restrictions on the conditions for the existence of the latter. Third, it allows us to distinguish in a clear way some given individuals from their corresponding set, as for the latter to exist, there must be one way or another by which they are brought together to form a Fine fusion, and indeed the identity of the set will depend on these various ways that they might be so related, whereras the identity of the individuals is fixed beforehand. Lastly, and related to the last point, we can now make sense of common claims by set theorists that the set of given individuals is just those individuals taken as one, and also Armstrong's view that there must be a "unit-making property" for there to be a set: here the difference between the present proposal and Caplan et al.'s view is simply that this property is not taken itself to be a distinguished formal part of a thereby also distinguished Fine fusion, but rather as a property of all sets, so that even by denying that sets are partially made up of the property *having some attribute or other*, or

[64] See Orilia and Paoletti (2020) for a discussion of different theories of properties in how they relate to the abundant/sparse distinction.

some similar state of affairs, we can still claim that all sets are such that their members are unified one way or the other, that is, by a given property or relation or other[65].

But, alas, equivalence classes are themselves *classes*, and thereby set theoretical constructions, so that the present view would be viciously circular, as sets are defined in terms of the notion of an equivalence class, which is itself defined as being a set, more precisely the set of objects such that they all bare R to each other, where R is a transitive, reflexive and symmetric relation. Still, thinking of sets as equivalence classes of Fine fusions has allowed us to highlight important benefits, corresponding to the points listed in the preceding paragraph, of what a solution sufficiently similar to it might provide. We now turn our attention to how a very similar view might be constructed.

Regardless of how close one might get to the notion of an equivalence class of Fine fusions, a good place to start is to define the appropriate equivalence relation, as follows:

Definition:

A given Fine fusion, x, y, z/R, is materially equivalent to another, u, v, w/R*, iff x = u, y = v and *z* = *w* or x= u, y =w, z = v, or x=v, y = u and z = w, or x = v, y = w and z = u, or x = w, y = u and z = v or x = w, y =v, and z =u[66].

After we're in possession of the relation of material equivalency, we can then define not an equivalence *class*, but perhaps something that can then play the same role and which we can take nonetheless to be the set containing the material parts of all of some given materially equivalent Fine fusions. Two hypotheses come immediately to mind at this point, when we consider the usual informal description of sets as forming a unity out of multiplicity: to take sets to be mereological sums of Fine fusions, and to take sets to be Fine fusions of Fine fusions. Here only the first view will be developed, as a full development of both would take too much space. It is left as work to be done in the future. Showing that one of the ways of thinking about sets is coherent and gives plausible answers to some difficulties pertaining to mereology and set theory is enough for our purposes.

First, let us consider the view according to which sets are mereological sums of materially-equivalent Fine fusions. The first obvious gain of this approach comes from Lewis's (1991, p.81) remarks about the ontological innocence of fusions: in a sense, if one accepts that given objects exist, then accepting them "taken as a whole" is not a further existential commitment, but rather just a different description of the same objects "taken separately"[67]. Set

---

[65] Crucially, here the property *being unified one way or the other* is not what sets generally *are* or a special part of what they are, but rather a property that all sets have. By making this move, we can also avoid several difficulties that have been raised by Paseau (2008) against the reduction of sets to other entities, as there are no objects to which the sets reduce to.

[66] Of course not all Fine fusions are defined on exactly three objects. Informally we can say that Fine fusions are materially equivalent if and only if they the same material parts. We could instead appeal to plural quantification and claim that x's/R and y's/R* are materially equivalent if and only if the x's are equal to the y's.

[67] See Bøhn (2011) for an interesting comment on these remarks by Lewis, supporting the present interpretation that Lewis *does* consider composition to be a form of identity, and where Lewis should (contrary to what he endorses in

theory would then itself be ontologically innocent, as given certain individuals and certain properties and relations that act as the formal parts of the given Fine fusions, nothing is added by accepting their fusion. It might seem strange that the set containing the ham and the two slices of bread as members would have the ham sandwich as a part, but if we consider sets to be ways of taking objects together as one without it mattering in which way in specific they are so, then we might say that the fusion of all the materially-equivalent Fine fusions to the ham sandwich are all such ways of taking the ham and the two slices of bread as one, and the ham sandwich is just one such way. Given that all fusions of Fine fusions will correspond to a set, we can then make more intuitive sense of the idea that the Fine fusions are part of the set, because the set has the property of unifying its members in some way or another, which might intuitively be grounded or explained by the fact that it contains all such ways of unifying the members (the Fine fusions) as parts.

The present view has as a consequence that whenever there is a Fine fusion, there also is a set corresponding to it. But is it the case that there are enough Fine fusions for all the sets of set theory? For instance, is it always the case that for any given collection of individuals, there is a relation that unifies them, so that there is a corresponding Fine fusion? I believe one can prove that there are in at least two ways: (i) if one accepts Lewis's claim that class inclusion is a mereological relation between classes and that every object must have a property; or (ii) if one accepts that different sets always have different properties. Here only the first method is developed, but to clarify how the second method would go, it can briefly be said that if different sets always have different properties, then it will be possible to always form a singleton from a given set, as the fusion which a given set is will then have different formal parts from those that the first set has, so that again a Fine fusion can be formed while respecting strong supplementation, and from that Fine fusion (and presumably others), one can get a further fusion of Fine fusions, which will be the singleton of the set that one has started with. In general, if different sets have different properties, and all individuals have singletons, then it will always be possible to form different sets from given sets. In what follows, we explore the first way of showing there are enough Fine fusions for all the sets, and touch on difficulties that both options face.

In the first way just mentioned and here developed, we start by proving that every object has a singleton, by assuming that every object has at least one property, showing that therefore there is one Fine fusion for every object, and therefore, given mereological universalism, that there is a fusion corresponding to the materially-equivalent Fine fusions whose material part is just the specified object. We then show that in general sets can be obtained from the singletons. Suppose that x and y are individuals, that x/P is a Fine fusion containing x as its material part and that y/F is a Fine fusion containing y as its material part. It follows, by the present view of sets, that x/P is part of {x} and y/F part of {y}. In order to get from {x} and {y} to their fusion, we fuse the material and formal parts of any of the Fine fusions of x with the material and formal parts of any of the Fine fusions of y, so that we get the Fine fusion x, y/(P or F). Here the fusion

other parts of the book) accept that some relations that seem to hold directly between certain objects, should actually be relativized to how certain objects are described - either as a single whole, or as multiple objects.

of properties is itself taken to be a disjunctive property, and we can think of this move as joining together the extensions of the original properties, so that a fusion of properties is just that property that applies to all the objects that instantiate the fused properties, as it contains both properties as proper parts.

To intuitively make sense of why we think of Fine fusions in this way, consider that in this ontologically innocent way of conceiving of wholes, for instance the fusion of a ham sandwich with a car is not some new bizarre object, but it's just both taken together, and if the relations that hold between their respective material parts are parts of it, then it seems that these relations will themselves also be part of the whole formed by their fusion, so that it seems that they should also be conjoined in the whole. One way to do so would be to take their conjunction, but this wouldn't work, as, say, the relation of *betweenness* won't apply to all the material parts of the new fusion, namely those material parts that are part of the car. It seems, then, that we should take the formal part of the new Fine fusion to just be the disjunction of the formal parts of the initial Fine fusions, as in that way we guarantee that the new formal part will apply to all the material parts. In fact, given how varied Fine fusions might be, and hence how unrelated, it seems that taking the disjunction of the initial formal parts is the only way to ensure that this is the case. After arriving at the Fine fusion x, y/(P or F) from taking one part of {x} and of {y} and fusing them together, we can then apply the same process for every other combination of parts of {x} and {y}, thereby getting all the Fine fusions of Fine fusions that have x or y as material parts. If we then fuse all these Fine fusions together, we get the set {x, y}. Since every object has a singleton and since it is possible to get all the other sets from the singletons, in a similar fashion to how Lewis does it, then it seems that all the sets can be obtained from this procedure[68].

Do we thereby have too many sets? Notably, Lewis had to accept the thesis that not every object has a singleton because of his commitment to unrestricted composition, for Lewis is committed to some classes - the proper classes - that are not members of any set, and so can't have singletons. By accepting mereological universalism, that is, that any objects, no matter how dispar, form a whole, and that every object has a singleton, the theory so far developed would run into the usual set-theoretical paradoxes. Either the assumption that any object has a singleton, or the assumption that any objects form a whole has to go. In the following paragraph I would like to argue that it is the assumption that any objects form a whole that should go, but that this does not affect what has been established thus far.

In order to justify the option for restricting composition, let us recall that one of Lewis's reasons to accept mereological universalism was that mereology is "ontologically innocent", that whenever we accept that given objects exist, it is no further ontological commitment to maintain that the whole formed by them also exists. As an example, Lewis gives the case of a six pack of

[68] A nice feature of the present proposal is how close the process of forming the fusion {x} and {y} is to the way the union set of two sets is formed. Just as in the case of the union set we have to consider what the members of the set are, and then unify them under a new set, in the case of the fusion of singletons we likewise consider the parts of each set and then combine them, fusing the resulting Fine fusions to form a new whole. In Lewis's view such a story is not available as the singletons are themselves mereologically atomic.

beer. If one goes to the supermarket, one does not buy each beer individually from the pack, and then pays an extra price for the whole; rather, one just pays for the six beers, whether they're described as separate beers, or as a pack. It would seem, then, that for Lewis's argument to be plausible, taking things to be a whole or to be separate is simply a matter of how we wish to describe the objects we're talking about - perhaps some descriptions are more useful, natural or recurrent, such as describing some atoms as a human, whereas other wholes are not useful descriptions, such as the trout-turkey. But it seems that's all there is to it. If this interpretation is right, then it is no wonder that philosophers such as Sider have moved from one "extreme" to another in the debate over what wholes there are, for given the ontologically innocent view of mereology, mereological universalism comes very close to mereological nihilism, the view that only mereological atoms exist, so that the extremes seem to almost touch.

It is not clear to me that there is a big gap between saying that there exists a table, as all of its proper parts exist and whenever they do, there's always a whole corresponding to it, while adding that accepting the whole is no further ontological commitment, and saying that there is no table, but only its proper parts that have themselves no proper parts, and that when we say true things about the table, that is on the account of a paraphrase where we speak not of a table, but of certain mereological atoms arranged in certain ways. If there is no further ontological commitment when we accept that there is a table after accepting that all of its proper parts exist, then likewise the same applies for the tables' first parts (as defined above), and second parts, and so on… If atomism is correct[69] and objects don't divide down to infinity, then it would seem that the mereological universalist would have to say that there is no new ontological commitment when one accepts the existence of the various wholes formed by the atoms of the table, just a different description of those same atoms either as a unit, or as scattered different objects. But if what it comes down to is just a difference in terms of description - "take them together or take them separately"[70] -, then it would seem that it is precisely that there is only a difference in terms of description that allows the nihilist, such as van Inwagen, to say that, for instance "There are atoms arranged table-wise" is a paraphrase for "There is a table".

Perhaps more importantly than its relation to mereological atomism, it seems that if mereology is indeed ontologically innocent, then Lewis has appealed to notions that represent in fact no gain compared to the usual informal claims of set theorists about what sets are. If taking things as a whole is just a different way of describing them, and the whole is nothing but its parts, then how is that different from claiming that a set is some objects taken as a unity, which Lewis criticizes: "To this day, when a student is first introduced to set theory, he is apt to be told something similar. 'A set is a collection of objects .... [It] is formed by gathering together certain objects to form a single object' (Shoenfield). A set or class is 'constituted by objects thought of together' (Kleene). 'Roughly speaking, a set is a collection of objects and is thought to have an

---

[69] The commitment to atomism seems to be, if these considerations are correct, to be the main difference between the two views. Mereological universalism does not seem to be committed to an atomist picture of reality, whereas if there are effectively no objects with proper parts, then the existence of gunk is ruled out, for if everything were to be made of atomless gunk, then everything would have proper parts, which is the direct opposite of the nihilist's claim.

[70] Lewis (1991, p.81).

independent existence of its own ... ' (Robbin). (...) But after a time, the unfortunate student is told that some classes - the singletons - have only a single member. Here is a just cause for student protest, if ever there was one. This time, he has no. 'many'. He has no elements or objects - I stress the plural- to be 'combined' or 'collected' or 'gathered together' into one, or to be 'thought of together as one'. Rather, he has just one single thing, the element, and he has another single thing, the singleton, and nothing he was told gives him the slightest guidance about what that one thing has to do with the other."[71].

Lewis's criticism of the preceding ways of trying to characterize sets might be legitimate, and as we have seen, the case of the singleton is a tough one to crack. Still, it would seem that his project of reducing set theory to mereology with plural quantification assumes that the notion of a mereological sum or fusion is better understood than talk of taking many objects to form one single object. But the notion of a sum has itself the same problem, for the sum, S, of any given two objects, x and y, is the whole whose proper parts are all parts of x or of y. In the case where x = y, and so when an object is summed to itself, then the resulting object is just x itself. But then a dissatisfied student could also complain that they do not understand how it is that an object is also the sum of itself (here notice that the sum of an object with itself won't be equal to two of the same object, for they are numerically the same, not qualitatively identical but distinct objects, case in which that would indeed correspond to their fusion, where "same" is qualitative sameness).

Of course, Lewis could then very well complain that while in the case of the singleton the set-theorist posits two different objects, in the case of mereology the sum of an object with itself is just that same object, so that no mysterious connection between distinct entities would arise. But even if so, while for all the other sums what sets the whole apart from the parts is simply that it corresponds to a different description of those same parts, in the case of the sum of an object with itself, there seems to be no difference in description. If there were, then it is not clear what the differing descriptions are. In the case where we only have one object we won't be able to make sense of the distinction between whole and part as in the case of the fusion of the cats where: "The fusion is nothing over and above the cats that compose it. It just is them. They just are it. Take them together or take them separately, the cats are the same portion of Reality either way. Commit yourself to their existence all together or one at a time, it's the same commitment either way. If you draw up an inventory of Reality according to your scheme of things, it would be double counting to list the cats and then also list their fusion."[72]. It seems that we cannot, then, make sense of the claim that every object is part of itself, for we have no grasp on what the differences are between taking a given object to be a whole, or to be a part of itself, an understanding that is required in all other cases, according to the universalist who also adheres to the ontological innocence of mereology.

Coming back to the question that led us to the preceding considerations about how the relation between a part and the whole it is a part of, from the perspective of someone like Lewis,

[71] Lewis (1991, p.29).

[72] Lewis (1991, p. 81).

might be similar to some ways of thinking of sets, we can then see, keeping up with the analogy, why we might want to restrict composition. If fusions are pretty much like sets in terms of taking multiple objects and taking them to form a unity, then perhaps the same limitations that can be imposed on the construction of sets also apply in the case of considering certain objects to form wholes. For instance, consider the case of the putative universal set, the set containing all sets, which is usually thought to not exist for reasons closely related to Russell's Paradox. In fact, if the preceding story of the metaphysics of sets is right, we can provide a non-arbitrary reason for why the universal set cannot exist only based on its supposed mereology: since the members of a given set are proper parts of it, the universal set would have to contain itself as a proper part, for it, being a set, would have to be a member of itself, violating the anti-symmetry of the relation of proper parthood[73]. We might then want to restrict on this basis set-membership, so that no set is ever a member of itself. In terms of the current picture, that would amount to the restriction that no fusion of materially equivalent Fine fusions ever has the exact same formal parts as the Fine fusions that are part of it, which amounts to the same as holding that no set instantiates all the same properties and relations as its members, which in particular also entails that no two sets ever have the same properties and are part of the same relations[74]. Here we take this claim, which is an instance of the principle of the indiscernibility of identicals, to be uncontroversial.

Let us try, then, to see what restrictions on composition there might be. From our assumption of classical mereology, it follows that no whole has itself as a proper part, so we can rule out compositions that violate anti-symmetry. But that is, of course, no restriction at all in the present case, for Lewis accepts unrestricted composition without accepting loops of proper parthood. What about the size of the wholes? It is often said that some collections are just too big to be sets. Can the same be said for some wholes? Are there some things such that they are too many to form a whole? Yes, if we accept, as we have, Fine fusions. Let us suppose that Reality is the mereological analogue of the universal set, that is, the whole that has everything as a part. Then Reality, as a whole, will have properties, for instance *being a whole*, *being infinite*, *being self-identical*, and so on. With Reality and all such properties, it is possible to form multiple Fine fusions, which are themselves objects. But that entails a contradiction, for then Reality would be a proper part of an object, whereas we have defined Reality to be the whole containing everything as a part. It seems, then, that composition should be restricted in order to disavow the existence of wholes that are, in a sense, too large.

[73] In this way, we can find some plausibility in the restriction of standard set theory, and in the theory of types, according to which no set is a member of itself, which is often thought to be an ad-hoc move, made to directly address the paradox whilst lacking independent motivation - for shouldn't for instance the set of numerous sets contain itself? In fact, given what has been said earlier, there are other reasons one might want to reject anti-symmetry. If the members of sets are proper parts of them, then cases like the set of numerous sets might provide further counterexamples to the anti-symmetry of the relation of proper parthood. This justification for the axiom of regularity is presented from the standpoint of a proponent of classical mereology, and so from the point of view of someone who accepts anti-symmetry. It is not necessarily endorsed. Later on anti-symmetry will be considered again in connection to the broader goal of the present section.

[74] This is exactly the assumption of the second way of thinking of sets in the present framework mentioned earlier. It may be that the two ways of accounting for sets turn out to be equivalent. That, however, is left for future developments.

Instead of showing the impossibility of Reality, one might take this *reductio* to show instead that one should limit Finean composition, that is, the formation of Fine fusions from given objects and properties. I believe, however, that even ignoring Fine fusions it will still be possible to show the impossibility of Reality. It seems that Reality should contain as a proper part all the properties and relations, for these are as well real, that is, they exist, whether they're taken to be tropes, immanent universals or anything else. But it would seem, as stated earlier, that Reality itself would have properties. But then it would seem that these properties would have to be, in a sense, outside of reality in order to apply to it as a whole. For instance if we accept that properties being instantiated by objects form facts, then it would seem that, if the property was wholly included in Reality, that the fact wouldn't express the connection between distinct objects, or between a property and an object, but would rather just correspond to Reality, as the relata are within the same object (as a proper part and a non-proper part). But Reality is itself, as it has been defined, not a fact, but rather an object. It seems, then, that the difficulties we are faced at the moment come rather from taking Reality to be a single object, that is, from assuming that there is a whole that contains everything as a part.

We haven't imposed any restrictions on Finean composition. Should we? If for every property and relation, plus some objects (except for the case of the empty set[75]), there is a Finean fusion, wouldn't we be able to form the set of all non-self members? Fortunately for the present view, the answer is that such a set cannot be constructed. In fact, the reason why the set of non-self membered sets could be formed in naïve set theory is that it contained an axiom of unrestricted comprehension. In the current setting, the axiom of unrestricted comprehension would amount to the claim that whenever there are some Fine fusions that share a formal part, we can fuse them together to get the corresponding set, which would supposedly be the set of objects which satisfy the relevant formal part, whatever they may be. We have not, however, committed ourselves to the view that those fusions correspond to sets, and Russell's paradox provides us with a reason to not take such fusions to be sets. In fact, it is easy to see why we don't need to take them to be, in general, sets, for while it is intuitive to take the fusion of materially equivalent Fine fusions to be sets, as they contain the same objects and in a certain sense make irrelevant what property or relation unifies them, as what counts is just that they are so unified, in this case we would have to say that it does not matter what objects a given property or relation applies to, just that it does apply to some objects. But to be irrelevant what objects a property or relation applies to does not mean it's indifferent how the objects are unified: they are unified in a specific way. Therefore to accept the principle of unrestricted comprehension would here correspond to a different intuitive way of thinking of sets, no longer in terms of certain objects being unified in some way or another by given properties and relations, but rather the other way around and sets would have as members whatever objects instantiate a given property

[75] Here we won't explore in detail the questions that arise when we consider the empty set. In the view I'm proposing, we can take the Fine fusions with no material part to be the Fine fusions such that no object satisfies their formal part. This does not commit us to the necessary existence of properties or a Platonist view of universals, as the formal parts can just be taken to be conjunctions of properties that no object satisfies - like *being round and square* - or formed out of properties that all objects instantiate - like *not being self-identical* - so that all uninstantiated properties depend for their existence on instantiated properties. The empty set can then be taken to be the fusion of all such Fine fusions.

or relation. We might, in this framework, then take naive set theory to correspond to the view which makes the assumption that whenever we consider whatever objects might be unified by a given property and relation, we might then take those objects and consider all the ways in which they may be unified, so that we can arrive at the corresponding set. In the paragraphs that follow we show why it is not possible to form the sets of all such objects, in virtue of the objects to which some properties, such as *being a non-self member* apply. Given that we might not know how the formal part behaves, we can't rule out inconsistencies - and it seems that it is precisely this that leads to cases as the putative set of non-self members. If intuitively we think of the material parts of the Fine fusions containing the same property or relation as a formal part as the extension of that property and relation, then effectively we are then capable of explaining why the extension of some properties and relations do not correspond to sets, as they do not form wholes, and therefore why a principle of unrestricted comprehension should not be accepted.

In the case explored in Russell's paradox, the property *being a non-self member* does apply to several objects, thereby forming Fine fusions, but it leads to contradiction if one takes the fusion of these Fine fusions to be the set of *all* non-self members, exactly because it is not irrelevant what objects the property applies to, as it can't apply to the fusion of Fine fusions who all have it as formal parts, which we can show simply based on mereology once again. In the present framework it will never be the case that a set figures as a material part of a Fine fusion that is then fused with other formally-equivalent Fine fusions to form that same set. The putative set of all non-self members, let us call it S, wouldn't contain itself on this view because it would be the fusion of Fine fusions containing as a formal part *being a non-self member*, and for it to be a part of the set, there would have to be a Fine fusion S/*being a non-self member*. But, given what has been said about what Fine fusions are like, S would be of the form [all the objects to which the formal part applies]/*being a non-self member*, so the relevant Fine fusion of that set would be [all the objects to which the formal part applies]/*being a non-self member/being a non-self member*. This Fine fusion, however, cannot exist in the framework here developed, as it would represent a violation of the principle of strong supplementation, just as Caplan et al.'s (2010) view according to which /*having some attribute or other* and /*having some attribute or other*/*having some attribute or other* were distinct Fine fusions, corresponding to the empty set and its singleton, whilst having the exact same parts.

The preceding considerations, however, show more than what was needed, for even if S is not considered to be a set, by strong supplementation we still have that there can't be a Fine fusion of the fusion of *all* the relevant formally-equivalent fusions. This fact could be taken to show that the argument against sets satisfying conditions such as *not being a self-member* overgenerates, and therefore that something must be wrong with it. On the contrary, however, it seems that there is already a problem with admitting all fusions of formally-equivalent Fine fusions, as - adhering to the principle that any objects instantiating a given property or relation form a Fine fusion - they will lead to contradiction. Take for instance the property *being self-identical*. Should we accept the fusion of all Fine fusions containing that property as its formal part? No, because effectively all objects instantiate that property, so that such a fusion would contain all of the objects as parts - it would correspond to Reality, which we have shown to be impossible. Cases like this and the preceding case of *being a non-self member* can be vastly multiplied - the only requirement needed to lead to contradiction is that both all of some formally-equivalent Fine fusions and their fusion instantiate the same property or relation. This illustrates what was said earlier that one can't in the same way take the objects that instantiate a property to be irrelevant as the formal part of a Fine fusion can in the case of sets.

This last conclusion does not entail that we should impose an arbitrary restriction on what fusions there are, for instance claiming that some formally-equivalent Fine fusions do not have a fusion. Rather, these examples show that we can't even begin to uphold the axiom of unrestricted comprehension, for the fusion of given Fine fusions that are formally-equivalent, when that fusion also instantiates the relevant formal part, won't be the case where it does not matter what objects instantiate the property (which would correspond to the axiom of unrestricted comprehension), for there is at least one object that is left out as a material part: the fusion itself. So we might say, in the end, that the reason why the axiom fails is that there are some formulae to which, by their nature and because of independently motivated restrictions on composition, namely the principle of strong supplementation, there does not correspond the fusion of all Fine fusions containing it as a formal part. There is a set containing some of the objects that satisfy the given formula, but by the very nature of the property (or properties and relations) involved, there is never a set of *all* such objects, for there is no Fine fusion corresponding to all of them. To say so might still seem to be arbitrary, but once one considers that in order for there to be such a fusion, all Fine fusions containing the relevant property as a formal part must exist, and that if all such Fine fusions existed, as per the nature of the property itself, some of them would have other such formally-equivalent Fine fusions as material parts, then, accepting strong supplementation, one has to reject one of these steps, then perhaps the most plausible option seems to be to reject that there are two Fine fusions with the exact same parts. If there is a restriction to Finean composition it is this. But, like before in the case of anti-symmetry, it is no restriction at all in the current setting, as classical mereology has been accepted from the get go.

The case of Fine fusions which are, in a sense, contained in themselves, are, however, but one case where it would seem that distinct Fine fusions have all the same material and formal parts. For instance Othello, Desdemona/*loves* and Desdemona, Othello/*loves* seem to correspond to different Fine fusions, and yet they still have the same material and formal parts. I believe that Fine is right in accepting such distinct Fine fusions, and that cases like this are further cases where strong supplementation fails. Still, since we're accepting that strong supplementation holds, we will ignore such Fine fusions in the current context: focusing only on the Fine fusions where the order of the material parts does not matter, we still have enough Fine fusions for all the sets, as we can get all the relevant Fine fusions starting from the properties that given individuals have, and then fusing the Fine fusions where the individuals figure as material parts. Much more would need to be said about structure, order and the principle of strong supplementation, but discussing them would take us too far - and we have already come very far from the original motivations of the present section. To ease a lingering feeling of arbitrary exclusion of certain Finean fusions, we might motivate this exclusion by considering that set theory is itself extensional. While the situation corresponding to Desdemona loving Othello and the situation of Othello loving Desdemona are distinct (and Fine (1999) individuates Fine fusions in terms of the individuation of situations), the set {Othello, Desdemona} is not sensitive to order, so that from the point of view of set theory sets are only distinguished by their members, and to identify the members, one does not need to consider cases where these are related in ways in which the relation has a direction. To put the point succinctly: given that set theory is extensional, it is natural to ensure that the underlying mereology likewise is extensional. Alternatively, one might impose the restriction that since all the necessary Fine fusions can be arrived at by taking fusions of Fine fusions which only contain properties and a unique object as a material part, then only these properties of objects and properties formed out of them should be allowed in the formation of further Fine fusions, thereby putting to the side Fine fusions to be formed where the formal parts have a direction, as it were, like the relation of *loving*. Again, much more would need to be said about this particular point, and a sense of arbitrariness remains, in spite of the extensional motivation.

It would seem, then, that taking sets to be fusions of materially-equivalent Fine fusions provides us with a plausible framework where multiple questions about the nature of sets and even about the justification of some famous restrictions of standard set theory are directly addressed. Taking sets to be such fusions, and fusions of given Fine fusions to be the Fine fusion containing all the material parts of the first Fine fusions and the disjunction of the formal parts of those same fusions, then intuitively we can think of sets as being all the ways of unifying certain objects, with the help of a property or relation. With this view we can also give an answer to one of our puzzlements: what distinguishes the singleton from its single member? The answer is that the latter is simply the individual itself, whereas the former is that individual taken in the various ways one can attribute properties to it, or relate it to itself (*being self-identical*, *having the same height as*[76], ...). The reason why the notion of a singleton is so puzzling is that it seems plausible that to be an individual is to be, following an Aristotelian tradition, a possible subject of predication, that is, for it to be something of which something can be said. Effectively, the individual remains the same, and perhaps to recognize it as such we must recognize it as something that instantiates properties and relations, and yet, there is a difference between taking an object as an individual, and taking it as an individual to which such-and-such properties apply. It might be, then, that to have an individual in mind is to therefore also have a grasp of its singleton. If a bizarre case can be constructed where an agent knows a given object exists without knowing what any of its properties are, then the agent could know that the individual exists, that the singleton exists (for they know that to each individual, there corresponds a singleton), without thereby knowing what the individual is like - what properties it instantiates -, and likewise not knowing what its singleton is. By taking the singleton of an object to be the aforementioned fusion, we can explain why it is the agent could know that the singleton exists - its existence does not depend on any of the properties of the object, but only on the existence of at least one property or relation that it instantiates - without knowing what the singleton is like - as knowledge of that implies at least some knowledge of what properties the object instantiates. Singletons and individuals are, then, still very closely related if an Aristotelian view of individuals is assumed, which helps explain the puzzlement that arises from the fact that they are nonetheless distinct. We can also then explain why for instance Quine took singletons and

[76] Assuming, of course, that the object is spatially extended.

individuals to be identical. If the present considerations are on the right track, then Quine was not far off the mark, for even if they're distinct, it might be a requirement that to have a grasp on the qualitative character of an individual, one has to have a limited grasp on the identity of its singleton.

Back to anti-symmetry?

In the previous section it was argued that someone that favours a linguistic ersatzist approach to possible worlds might want to maintain that the relation of proper parthood is not antisymmetric. In this section we have, however, on our quest to commit as little as possible to controversial metaphysical claims, and assuming, to give more credibility to arguments against what has been so far maintained, that classical mereology is still the most uncontroversial theory of part and whole on offer, reached the conclusion that no set is a member of itself, which is in stark contrast to what was maintained in the last section, where we thought plausible that a given possible world, which is a set, contains itself as a proper part.

One other way in which the approaches differ is in that while here we have thought of sets as fusions of Fine fusions, and the members of sets to be individuals or Fine fusions, the sets that correspond to possible worlds have as elements sentences from the lagadonian language, which are ordered tuples of individuals, relations, properties and set-theoretical constructions, rather than individuals themselves, or even Fine fusions. Going back to a previous example, one thing is the ham sandwich, the other is the ordered tuple containing the two slices of bread, the ham, and the *betweenness* relation. But, Fine fusions or not, lagadonian language sentences can still figure as the material part of Fine fusions, so that it is possible to form sets out of the lagadonian language sentences. There is, then, no obstacle arising from this distinction to taking possible worlds, as described earlier, to be sets, where sets are understood as in the preceding pages. There is, nonetheless, an important aspect stemming from this distinction: the objects that figure in the lagadonian language sentences are not themselves members of the set that corresponds to the possible world referred to earlier. This aligns with Wittgenstein's (1921) claim that the world is the totality of facts, not of things (and likewise for every other possible world, with the difference that they represent states of affairs that do not obtain, whereas facts are obtaining states of affairs, in Wittgenstein's terminology).

How should we resolve this tension between giving up anti-symmetry, or find some other way of tackling the problem of alien objects and properties, on the one hand, and retain the view of the nature of sets exposed in this section? In the following paragraphs I try to give the argument against anti-symmetry more plausibility by showing that there is no such conflict.

The way here explored of dealing with the conflicting goals we have had in mind is to give up on trying to avoid all controversial metaphysical claims, provided we have avoided all unnecessary such claims, and accept the commitment to a non-antisymmetric relation of proper parthood. The task would then be to change the metaphysics of sets so far presented, to allow for cases where anti-symmetry does not hold. It would seem, however, that it was on the back of anti-symmetry that we have rejected some unwelcome sets, such as the universal set, so it would seem that by giving up anti-symmetry we would therefore lack the resources to do so. That is not the case. While anti-symmetry provides us with an immediate way of ruling out self-membered sets, the present framework can by itself show that no such sets exist. To see why, suppose that everything is as before, except that anti-symmetry does not hold. Suppose further, to make the case simpler, that x corresponds to {x, y}, where y is an individual. Let us, finally, suppose that F, F', F'', … correspond to all properties instantiated by x, that G, G', G'', … correspond to all properties instantiated by y, and that R, R', R'', … correspond to all relations holding between x and y. The set {x, y} would then be (x,y/FvG + x, y/F'vG + x,y/FvG' + … + x,y/R + x,y/R' + …). But crucially all the fusions of the form x,y/P where P is not a relation holding between x and y are themselves fusions of Fine fusions containing x and y individually as material parts, that is, of the fusions x/F, x/F'... and y/G, y/G'... . Since the order in which objects are fused does not matter, we can group all of the fusions containing only x as a material part and take their fusion, which is just {x} and fuse all the other Fine fusions that belong to the set {x, y} in what we might call *block*, to simplify. We would then have that, if x = {x,y}, then x = {x} + *block*. This entails that {x} is a proper part of x, which is not itself surprising since x is itself a proper part of {x} and proper parthood is not being assumed to be antisymmetric. In cases where a set is a member of itself, however, the set and one of its members all have the same properties, for they are the same object. But that means that the set formed by x itself and y will have all of the same properties as x does, so that they have exactly the same Fine fusions. As we saw earlier, no distinct Fine fusions can have the exact same formal and material parts, but if x = {x,y} then the Fine fusions of the set will be the same as the Fine fusions of one of its members, but if x, taken as the set of itself and y, cannot have distinct Fine fusions from those of one of its members, then it won't be possible to form a set from said Fine fusions, which both means that x would be a set without a singleton, contrary to our previous claim that every object has a singleton, and that it shouldn't be possible to form the set in the first place. It would seem, then, that by strong supplementation alone we can show that there is no such set, as such a case would require that two sets can be formed out of the fusion of the same Fine fusions: {x} and {x,y}, which is just x. Even if such sets were allowed and the preceding argument is not sound, however, it would not follow that the universal set exists, or that we would face Russell's Paradox, for in the present framework, by strong supplementation, it is still the case that we can't uphold a principle of unrestricted comprehension, so that we can't form the sets whose members are all the objects satisfying properties such as *being a non-self member* or *having all the sets as members*. And perhaps we could then allow for such putative sets as the set of numerous sets, which presumably should include itself, without thereby falling prey to Russell's paradox, avoiding at the same time what can be thought of to be arbitrary restrictions on sets, and the usual paradoxes that come without such restrictions.

But, given that we have said that we should give up anti-symmetry, wouldn't this last result make the case against such a position even stronger? After all, we just tried to show that even if we reject anti-symmetry, we can still show that sets cannot have themselves as members, if we think of sets as fusions of Fine fusions containing their members as material parts.

Here it is important to note an important distinction that has been stressed before: in cases where worlds contain lagadonian language sentences that reference the worlds they are part of, it is not the case that the world in question is a *member* of itself, but rather only, by the transitivity of the relation of proper parthood, that it is a *proper part* of itself. And importantly enough, even if the preceding argument against self-membered sets is right, there is no similar argument to show that accepting loops of proper parthood leads to contradiction. Let x be the set formed by the fusion of the Fine fusions that have y as a material part, and also one of its proper parts that is not a member of x. Given that x is a part of y (as y is the only member of x and we have seen that x is a proper part of x that is not a member of x, and is likewise not a property or relation), there seem to be only two options for what y is: either a mereological fusion of x and one or more objects; or a Fine fusion where x would at least partially constitute its material part. In the first case, the Fine fusions that x is a fusion of are of the form (x + z + …)/P, which is not the same as being of the form x,z,.../P - the former contains only one material part and the latter contains at least two. Further, it is not in general the case that if there is a Fine fusion (x + z + …)/P, that there is a Fine fusion x/P - to see why consider cases in which the fusion has properties that are not had by its proper parts, for instance the property *being a set*, which is a property of the fusion of given materially-equivalent Fine fusions, but not of any of them in particular. Since x/P does not in general exist on this basis, then it does not seem that the earlier argument applies, as the formal parts of the Fine fusions that constitute x won't be the same as those of the Fine fusions in which x is a material part. In the second case y will itself be a Fine fusion whose material part includes x, and it would then be of the form (x, z, …)/P. But, x itself being a set, y will be of the form (supposing x only has one member) (w/F + w/F' + …)/P. And again from here one can't show that there would have to be different Fine fusions containing the same material and formal parts, so one cannot conclude on the basis of the principle of strong supplementation that there can't be a set that has itself as a proper part.

In fact, there is a stronger reply one can give in this context, considering the discussion at the beginning of the present section in regards to what the parts of sets are - whether they include all the parts of their members, or are just given in terms of the ancestral of the membership relation. If one takes a Finean approach here, then y's parts are not themselves parts of x, so that we might even adopt anti-symmetry, but still deny that the proper parts of members are themselves parts of the sets of which they're members, so that there would be no objection to taking a given possible world to also figure in some of the lagadonian language sentences it contains as members. If the preceding considerations are right, however, then even if one adopts a view of the mereology of sets according to which the parts of members of sets are also parts of sets, we can retain the metaphysics of sets so far explored and still give up anti-symmetry, which strengthens the case for rejecting it[77].

[77] A different reason one might think that even if the other principles of classical mereology should be accepted, anti-symmetry should be independently rejected is that anti-symmetry *is* independent of the other principles. For instance Fine (2010, p. 568) shows that transitivity and reflexivity hold simply in virtue of the definition of the relation of parthood in his operational mereology, but that anti-symmetry only holds if one assumes anti-cyclicity, that is, that there are no loops in the chain of applications of a given operator, which claims that if x can be composed from itself, then every intermediate whole between x and x should itself be x. However, to accept anti-cyclicity amounts to the same as directly rejecting anti-symmetry, for it is only in cases where there are loops of proper parthood relations that there is anti-symmetry. It seems, then, that to avoid anti-symmetry in this way is an *ad-hoc* move which requires further argument. Cotnoir and Bacon (2012) show how antisymmetry is an axiom of classical extensional mereology, and how it is not derived from the other principles of classical mereology. It would

Sentences can refer to themselves, if the actual world corresponds to the totality of facts then it includes facts about itself, and perhaps some propositions and time travelling objects make up likewise proper parts of themselves. I take that in at least some of these cases the appearances do not illude, that anti-symmetry is in fact violated by some wholes and their parts. I take it, therefore, that we should not coil and try to avoid counterexamples to anti-symmetry, but accept that the relation of proper parthood *does* in some cases lead to loops. If there is a controversial metaphysical claim that the perspective here developed is committed to, it is this.

Summary

In order to avoid the difficulties that plague magical ersatzism, we reached the conclusion that worlds should be thought of as set theoretical constructions whose ultimate constituents are concrete. At the same time, however, it seemed like there were a number of difficult issues surrounding the claim that sets contain their members as parts, starting from a putative conflict between the transitivity of the relation of parthood and the fact that set-membership is not in general transitive. Using resources from Finean mereology, while rejecting some of its non-classical commitments, we have been able to provide an account of the nature of sets according to which the members of sets are parts of sets in a way that meets challenges that alternative views face, such as: rejecting the principle of strong supplementation; and taking the relation between a given object and its singleton to be a mysterious and yet necessary primitive. We can, then, meet Lewis's objection and claim that possible worlds are sets of lagadonian language sentences, which, being members of worlds, are parts of it, where these sentences can themselves taken to be concrete or to contain concrete proper parts, such as individuals and tropes or immanent universals.

Finally, we considered a potential tension between the views for how to address the problem of alien objects and properties, and the metaphysics of sets just proposed, showing how such a tension is merely apparent as nothing in this metaphysics of sets turns on the acceptance of anti-symmetry as a general condition that the relation of proper parthood must meet.

---

seem, then, that our treatment of anti-symmetry as being in some way separate from the other principles of classical mereology tracks the way in which those principles can be established, which, to some extent, makes the move seem less gratuitous.

From the possible to the impossible

Having developed a linguistic ersatzist account of the nature of possible worlds and replied to objections to it, we can now turn to the question of how to expand the present view to accommodate impossible worlds. Given the objections so far considered for the other versions of ersatzism and given that accounts of possible worlds as genuine can't be extended to accommodate impossible worlds, a metaphysically satisfactory answer to the problem of logical omniscience that relies on enriching the metaphysical framework of possible worlds with impossible worlds depends on whether linguistic ersatzism (in how it has been so far developed) is compatible with this expansion.

An obvious way of accounting for impossible worlds in the linguistic ersatzism framework so far explored is to take impossible worlds to be all the sets of sentences of the lagadonian language that are not both maximal and consistent, so that the possible worlds would correspond to the maximally consistent sets of sentences, and the impossible worlds to all the others. This way of defining "impossible world" has the consequence that impossible worlds are very anarchic, in the sense that there seem to be little to no rules on what a given world has to represent, given what it represents. In fact, since for every sentence there will be a world that represents it and nothing else, worlds are only "constrained" by the trivial rule of identity, that is, the rule according to which if a given world represents that A, where A is any given sentence of the lagadonian language, then that world represents that A. Worlds construed in this way have been called "open worlds" by Priest (2005), in the intuitive sense that it is totally open what they represent, as there are no non-trivial logical constraints on what they represent. By contrast, in the case of possible worlds, representation is closed under (let us assume) classical logical consequence, so that there are strong constraints on what they represent.

The decision to count all the sets of sentences of the lagadonian language as worlds is controversial. On the one hand, one might think that impossible worlds ought to be in some sense complete or maximal, and therefore that we should only drop the consistency requirement imposed on possible worlds: "incomplete worlds" that are not inconsistent are not, in any respectable sense, impossible, or even worlds - they are just situations or states, as described for instance in situation semantics[78]. On the other hand, one might think that impossible worlds should not be as anarchic as open worlds are, that impossible worlds should simply be closed under a logic weaker than classical logic, for instance a paraconsistent logic.

In what follows it will be assumed that impossible worlds are open worlds, and therefore reasons must be presented for why we should take every set of sentences of the lagadonian language to constitute a world. The first worry - that incomplete but consistent situations are not (impossible) worlds - can be met by considering the intuitive notion of a possible world, and the corresponding intuitive notion of what an impossible world is. As Nolan (1997) puts it, just as possible worlds can be thought of as ways that the world could have been, impossible worlds

[78] See Kratzer (2019) for an overview of the work done in situation semantics.

correspond to ways that the world could *not* have been[79]. It seems, further, that the world could not have been incomplete. For instance, it does not seem that the world could have been in such a way that I ate breakfast, but there is no fact of the matter (independently of any evidence that agents might have in that world) as to what in specific I had for breakfast[80]. As for the second worry, it seems that how much structure we should impose on worlds depends on what work we want the relevant impossible worlds to do. Our present considerations are aimed at providing a solution for the problem of logical omniscience, and hence the worlds needed must be pretty anarchic, as human-like agents seem to, at times, believe for instance the premises of a *modus ponens* inference, but not its conclusion, or to believe that *p* but not *p or q*, the same applying for other reasoning steps often taken to be trivial[81]. A different but related goal is to account for the nature of impossible worlds, regardless of the specific theoretical applications that they might have, so that an account that is suitable for every such application must be provided[82]. Some such uses, as we have seen should be the case for the problem of logical omniscience, require that some impossible worlds not be closed under any relation of logical consequence[83]. We should,

---

[79] It is disputable that the argument from ways for the existence of possible worlds is successful, and further that the same argument can be presented for the existence of impossible worlds (even if it could, that would certainly not be the case if we take worlds to be genuine, as Lewis, who proposed the argument from ways, does). The argument for the existence of impossible worlds from the existence of ways the world could not be, however, is not being presupposed here. Rather, there is merely an appeal to the intuitive notions of a possible and impossible world. Independent arguments for the existence of impossible worlds have and will be considered. The point presented here could alternatively be made appealing instead to Nolan's Principle (NP), which states that if *p* is impossible, then there is an impossible world that represents that *p*. It would be impossible for an incomplete situation to be a full description of what the world is like, so, by (NP), there is an impossible world which represents only what that situation represents (more precisely, this conclusion would have to proceed rather from the fact that an incomplete description, when taken by itself to be a world, is impossible).

[80] Here the law of excluded middle is not being presupposed to hold at all possible worlds. A world where a given proposition, *p*, is indeterminate, for instance, such that *p or ~p* is not true in that world is a world that represents *p* as indeterminate, and not a world that does not represent anything in regards to *p*, it is a world that does represent something in regards to *p* - namely that it is indeterminate. The present discussion, then, is orthogonal to questions about vagueness. Incomplete worlds in regards to *p* are, rather, worlds that do not belong to *p* (in the view of propositions as sets of worlds), or worlds that do not have *p* as a member (in the view of worlds as sets of propositions). Therefore, incomplete worlds are worlds in which the law of excluded middle fails (which might not be the case at possible worlds, depending on what one takes the correct position to be in regards to questions of vagueness).

[81] See Jago (2014) for a solution to the problem of logical omniscience that relies on the claim that what reasoning steps count as trivial is vague, and that "epistemic oversights" - that is, failures on the part of agents to draw a trivial conclusion from a given set of known premises - cannot be attributed to minimally rational agents. In the next section I will come back to Jago's position and propose an alternative to it.

[82] Priest (forthcoming) endorses this same goal.

[83] As it shall be seen in the next section and in the second part of this dissertation, this is not to say that no closure principles should hold on epistemically possible worlds. Hawke et al. (2020) argue, plausibly, that human-like epistemic agents knowing a given conjunction involves their knowing the conjuncts, so that no agent who knows a given conjunction fails to know any of its conjuncts. This principle will be considered later on. For now, however, the claim is simply that impossible worlds should not be closed under any relation of logical consequence, even if weaker than classical logic, as that entails logical omniscience in respect to the chosen logic, which bounded reasoners, like the average human being, are not. Further, even if the epistemic agents we wished to describe were, in regards to a weaker-than-classical logic, omniscient, we might have been interested in describing more limited

then, allow for worlds that have little to no constraints on what they represent. Finally, by allowing impossible worlds to be open worlds, we guarantee that we have enough worlds for our needs and whenever only a subclass of them are relevant for a given purpose or in a given discussion, we can then simply direct our attention to that subclass of all the worlds. In the same way that we might restrict our attention to the metaphysically or logically possible worlds, we might for instance want to restrict our attention to the epistemically possible worlds, which might comprise possible, as well as impossible, worlds, albeit not all of them.

Here an objection that might be raised is that since we accept Nolan's Principle (that is, a principle according to which for every impossible proposition there is an impossible world that represents it), even the rule of identity would be violated. It is impossible for a world to represent that A and at the same time not represent that A. Since this is an impossibility, by Nolan's Principle we should accept that there is an impossible world that represents that A, but does not represent that A. This however, we can't do in the framework so far presented, as even open worlds respect the trivial representation rule of identity. To accept such a world, further, would be to endorse a contradiction, just as much as we would be accepting contradictions in extended modal realism, that is, the view that there are possible and impossible worlds, all of which are genuine[84].

This objection is, however, misguided. While it is true that it is impossible that a world represents that A and does not represent that A, what this impossibility entails, by Nolan's Principle, is simply that there is an impossible world which represents that there is a world that represents that A, but does not represent that A. This however, does not commit the friend of impossible worlds to contradictions, as on the current proposal, there are no true contradictions *simpliciter*, but simply worlds that may represent contradictions. Impossible worlds that represent contradictions by containing a given sentence of the lagadonian language and its negation are what we might call blatantly inconsistent and they may not be epistemically possible for any agent (Jago (2014), Berto and Jago (2019)), but they are still, by the definition given earlier - where any set of sentences corresponds to a world - impossible worlds. By Nolan's Principle it is simply the case that for every impossibility there is an impossible world representing it, so that for instance the impossibility of there being a world, w, representing A but not representing A simply implies that there is an impossible world, w', according to which w represents that A and also that w does not represent that A. It is w', not w, which represents a contradiction, but since w' is an impossible world, there is no commitment to the truth of a contradiction, only to the existence of a world representing a contradiction.

A different objection consists in claiming that some propositions will not come out true in any world, if the present framework is to be able to provide a compositional theory of meaning. The objection is put forward by Stalnaker (1996), and can be presented by ways of a dilemma.

agents, or even agents that fail to be minimally logically competent, and in those cases nothing short of the complete anarchy of open worlds would serve to do the job.

[84] See Yagisawa (2010) for a defense of extended modal realism. This position has been met with strong objections by Berto and Jago (2019). For this reason, I don't have much to add to this particular discussion here.

Consider any two propositions, *p* and *q*, and their respective conjunction, *p&q*. It would seem that the meaning of the conjunction is a function of the meaning of its conjuncts, and of the truth-functional connective, “&”, and, assuming compositionality, that we should be able to retrieve the semantic value of “p&q” from the semantic values of “p”, “q” and “&”. If we also take the semantic values of propositional symbols to just be truth conditions, identifying propositions with sets of worlds, we can then present a dilemma for the friend of impossible worlds: is there an impossible world that represents that *p*, that *q*, but not that *p&q*? If so, then it seems that the semantic value of *p&q* is not, after all, determined by the semantic values of its conjuncts, for there is a world in which both conjuncts are true, but in which the conjunction is not so, so that the set of worlds that *p&q* is is not determined by the sets of worlds that *p* and *q* are, namely, by taking their intersection. If, on the other hand, there is no such impossible world, then it seems that there is an impossibility to which there corresponds no impossible world, contrary to (NP)[85]. If we give up (NP), however, then it becomes mysterious what criteria we should use in order to determine whether a given impossibility corresponds to an impossible world or not, and therefore we stop having a guarantee that there is an impossible world for every impossibility that we need an impossible world to correspond to.

One of the main reasons to accept impossible worlds is that they allow us to distinguish between necessarily equivalent propositions. The way they allow us to do so also displays one of this solution’s main advantages: leaving for the most part intact the possible worlds framework, which has been widely applied both in Philosophy and other academic areas. In an impossible worlds’ framework, while necessarily equivalent propositions are true in the same possible worlds, if they’re not the same proposition, they will differ in terms of what impossible worlds they are true in. Everything else remains the same, with propositions still being conceived as sets of worlds (now impossible as well as possible), the tool of accessibility relations on worlds can still be used to categorize certain interesting notions, as well as various other features of standard possible worlds models[86]. Let us assume again that possible worlds are closed under classical

---

[85] Stalnaker does not present the objection as it is presented here, but, I take it, the tension can be made clearer if we express it in terms of accepting or rejecting (NP), for it seems that the second horn of the dilemma is only pressing if one accepts that a friend of impossible worlds should endorse (NP).

[86] The main difference between possible worlds frameworks and ones expanded to include impossible worlds pertains to how the valuation function behaves. In impossible worlds a given proposition can be both true and false, so that we don’t have a valuation *function*, but rather a relation, taking at each world a given proposition to True, False, both, or neither (case in which the valuation relation does not hold in a given world for a given proposition and the world is thereby incomplete, as it fails to represent anything whatsoever in regards to the specified proposition). Possible worlds would then be the limit case where the valuation relation returns only one value for each proposition, and where the relation holds for every proposition. Further, in impossible worlds the truth value(s) or lack thereof of a given proposition are arbitrarily assigned, even for non-atomic propositions, whereas in possible worlds the truth-values of all such propositions are fully determined by the truth-values of the atomic propositions. This affords the impossible worlds framework much greater flexibility in the models that can be constructed from it. Timothy Williamson has claimed that this gain in flexibility entails a great loss in terms of explanatory power that is not compensated by other theoretical benefits. But Williamson’s objection is not so different, all things considered, from the objection considered here that has been put forth by Stalnaker, for it is precisely this arbitrary attribution of truth-values to propositions that both grants the impossible world theorist a greater flexibility in designing their models, and has the consequence that the truth-value of, for instance, a conjunction at a world will not depend on the truth values of its conjuncts. If the friend of impossible worlds can provide a different way of explaining how to

logic and let us consider an arbitrary proposition, *p*, and its logical equivalent *p v (p&q)* where *q* is again an arbitrary proposition but distinct from *p*. Let us refer, for ease of expression, to *p v (p&q)* just as *r*. Intuitively, *p* and *r* are true in the same logically possible worlds, yet, they are distinct propositions, and one but not the other may even be believed by an epistemic agent. Since this instance of necessary equivalence does not depend on the content of any proposition, it seems that using this method we can always find, for a given proposition, a necessarily equivalent proposition that we would like to distinguish from it. The reason why *p* and *r* are necessarily equivalent is that they're both true in all the *p*-possible worlds and false in all the other possible worlds, and this is so, intuitively, because *p* is just *p*, and because of how *r* is constructed, as it renders the truth-value of *q* irrelevant for the determination of the truth-value of the proposition it is a part of - the truth-value of *p* suffices to determine the truth-value of *r*. But just like in the case of *p&q*, once we move to impossible worlds the truth-value of *p* won't suffice to establish the truth-value of *r*, as we won't be able to determine the set of *r*-worlds from the set of *p*-worlds, for the former won't be identical to, or constructible from, the latter. It would seem, then, that for every proposition we could provide an instance of Stalnaker's objection from compositionality, so that instead of having no guarantee that there is an impossible world for every distinction between necessarily equivalent propositions one would like to make, it would seem we would now have the guarantee that there would *not* be enough impossible worlds, if compositionality is to be maintained.

While, as we saw, the current proposal is to be independent of any application that impossible worlds might have - so that the solutions for the objections to impossible worlds should also be so independent -, it is noteworthy that even if we were to accept only as fine grained impossible worlds as the solution for the problem of logical omniscience and for a correct characterization of epistemic possibility would require, we would still have to face the case just presented. Interestingly, however, that does not seem to be the case for the failure of closure under conjunction. It is very plausible indeed that sometimes epistemic agents fail to put together the information they have available so that they might believe that *p*, that *q*, but fail to believe that *p&q*. Lewis (1982) presents a case of this phenomenon: apparently he believed that Nassau Street ran roughly North-South, that the train station nearby was oriented roughly East-West but also that the two were roughly parallel. Far from being a controversial description of an unusual event, what Lewis described seems to be a common occurrence in human daily lives as agents that are presented and must integrate great quantities of diverse information. It would seem, then, that an impossible worlds solution for the problem of logical omniscience should either accept a fragmentist story of how human-like agents store information and how they reason based on that information, or accept that some impossible worlds represent *p*, *q*, but not *p&q*. It is not clear to me that they could not accept that a limited agent's belief system is

obtain the truth-conditions of complex formulae from that of its constituents, then it is not clear that the gain in flexibility compromises the theory's explanatory power.

fragmented, aside from admitting impossible worlds, as it seems that some of the motivations for accepting fragmentation are independent of the motivations for accepting impossible worlds[87].

The case against less fine-grained impossible worlds for epistemic possibility on the basis of the case of the failure of closure under conjunction does not seem, then, very strong. However, we face the objection from compositionality in its full force if we focus on the second case considered above. Suppose that *p* is the proposition expressed by "England will go to war with France" and *q* is "A nuclear bomb will be launched", so that *r* is "Either England will go to war with France or England will go to war with France and a nuclear bomb will be launched". As seen before, *r* is true in all and only the *p*-worlds, so that the truth-value of *p* determines the truth-value of *r*, independently of the value that *q* might take, so that if we want the compositionality facts to hold at all worlds, possible or impossible, we need *p* and *r* to correspond to the same set of worlds. But agents might believe or know that *p* and fail to believe or know that *r*, for instance if they lack the concept NUCLEAR BOMB, close to Stalnaker's (1984) example where King William III is wondering whether a war will take place, but nuclear military technology was yet to be discovered, so that he lacks the relevant concept and therefore is not wondering whether a nuclear war will take place. Impossible worlds that maintain compositionality fail, then, to distinguish between necessarily equivalent propositions that agents might have differing attitudes towards, but this was exactly the reason to accept impossible worlds in the first place: to disentangle necessarily equivalent propositions, or at least those propositions that agents might have different attitudes towards. Adding impossible worlds that respect all the compositionality facts would be a superfluous move that would not do the very work they would be added to do. An impossible worlds solution for the problem of logical omniscience would then necessarily violate compositionality, and, given the essential role played by compositionality in theories of meaning, it should then be rejected. In fact, if this objection is right, no theory that simply adds more worlds to possible worlds frameworks would succeed as either necessarily equivalent propositions such as *p* and *r* correspond to the same set of worlds and the theory gets the compositionality facts right, but then the addition is superfluous since *p* and *r* still cannot be distinguished; or necessarily equivalent propositions such as *p* and *r* correspond to different sets of worlds, but then compositionality ceases to hold.

In the preceding paragraphs I tried to present the compositionality objection to the impossible worlds solution in what I take to be its strongest version, and in the best light. As it is, it's a strong objection against accepting impossible worlds in one's ontology, and in specific for the purpose of solving the problem of logical omniscience. I believe, however, that it can be met, if one gives up a crucial assumption of the argument as it has been so far presented: that to accommodate the compositionality facts, a given complex proposition's truth-value in a given world should depend on the truth-values of the atomic propositions that compose it *in that same world*. That is, the assumption that the way in which the semantic value of a sentence depends on the semantic value of another should be internal to a given world, that the value of for instance "p&q" in a given world will be dependent on the values of "p" and "q" *in that same world*. It is

[87] Yalcin (forthcoming) discusses some reasons for why one might want to say that agents are fragmented, even if they're not cognitively limited.

easy to see that compositionality, so conceived, is at odds with impossible worlds, as it is precisely how formulas are valuated at worlds that is primarily changed in the context of an expansion of the possible worlds framework to accommodate impossible worlds, where complex formulas are assigned arbitrary truth-values that are in no way dependent on the truth-values of the atomic formulas (which might not even have a value at some impossible worlds). Given the immediate conflict between this notion of compositionality and the expansion envisaged by the friend of impossible worlds, they might argue that we should look for a different way of maintaining compositionality, one where its essential features and its important role in the explanation of how agents are able to learn languages and new forms of expression in a given language are still present.

Of course, it is not a general requirement of compositionality that the semantic values of certain formulae must be determined from the semantic values of other formulae or its components at the same world, so for instance the modal operator 'Necessarily' when added to a formula yields a formula whose semantic value is not determined by the semantic value of the formula it applies to in that same world, or at least only by that value, but rather by the values the formula takes at all accessible worlds. The requirement can be read instead as applying to non-modal formulae. This way of violating the intra-world requirement in the for non-modal formulae is also violated for instance by the so called Australian Plan for negation, according to which, as Berto and Restall (2019, p. 1121) put it: "whether $\neg C$ holds *here* depends on whether C holds *elsewhere*, in the same way that whether $\Box C$ holds in *this* possible world depends on whether C holds in other (*relatively possible*) worlds.". As we will see, the current way in which this requirement is met is not given in terms of what holds in other valuation points (the worlds), but rather it is given in terms of the structure of the entities that play the role of valuation points.

As a starting point, we might take compositionality to just be the feature of a language whereby from the semantic values of the linguistic constituents of a given complex term, phrase or sentence we can determine the semantic value of that term, phrase or sentence. The conception of compositionality just explored adds to this general conception two assumptions: i) that the semantic value of sentences are propositions, which are in turn truth-conditions, that is, the set of worlds in which the proposition expressed by the sentence is true; and ii) that the truth-value that a given non-modal formula has at a world should be determined in terms of the truth-values that its constituents have at that world, which are in turn determined from their semantic value.

The first assumption has deep historical roots in the analytic tradition, at least since the Tractatus, where Wittgenstein maintained that to know a proposition is to know under which conditions it is true. It is also an assumption shared by what could be taken as the standard view of the nature of propositions, the view according to which they are just sets of possible worlds, each of which can be thought of as a condition under which the proposition in question is true. This assumption won't be challenged or discussed for now but there are two points worth stressing in regards to it.

The first point pertains to how to characterize the assumption in a framework that has been expanded with impossible worlds. While possible worlds can perhaps be intuitively thought of as different conditions under which the propositions to which they belong are true, it might be less clear that impossible worlds can be thought of in this way, for since they are not ways that the world could have been, it might seem that they do not correspond to conditions under which given propositions would be true, as it is not clear how "would" should be interpreted here, given that since no impossible world could have been the actual world, the strategy of appealing to what would be true, had a given impossible world been the actual world, is not available. But while the actual world could not have been one of the impossible worlds, that only means, for an impossible world, *w*, and a given proposition, *p*, that "had *w* been the actual world, *p* would be true" is a counterpossible statement, that is, a counterfactual with an impossible antecedent. In an impossible worlds framework, counterpossibles do not come out vacuously true, so that it would seem that worlds can still be thought of as different conditions under which given propositions are true. Alternatively, one might consider impossible worlds to be conceivable, and take the antecedent (that is, that a given impossible world is the actual world) to be a mere assumption for the acts of imagining or conceiving that the agent, considering what would be true if a given situation held, performs. Regardless of the position taken in regards to how to conceive of impossible worlds as conditions under which given propositions are true, it is important to note that they do not represent a further commitment of the impossible worlds approach. Cases like Goldbach's Conjecture, previously used to motivate accepting impossible worlds in one's ontology, already require that agents are able to conceive of impossibilities (for either the conjecture or its negation is impossible, and it seems that both can be entertained by agents), and therefore also require that agents should conceive of some counterpossible statements that have them as antecedents, for instance: "had Goldbach's Conjecture been false, then a lot of Mathematicians' efforts would be frustrated", assuming that Goldbach's Conjecture is actually (and necessarily) true. It would seem then that taking a sentence's semantic value to be its truth-conditions is not at odds with adding impossible worlds to the usual possible worlds framework.

The second point is that while it has been relatively uncontested that the semantic values of sentences include their truth-conditions, there has been a lively debate in recent years, on whether sentences' semantic value is reducible to truth-conditions. It has been proposed, notably by Yablo (2014), that besides in what worlds a proposition is true, it is also part of the semantic value of a sentence *how* it is true at the worlds in which it is true, which corresponds to the sentence's subject matter or what it is about. The resulting view, which we can call two-component (2C) semantics, following Berto et al. (unpublished) can be used to try to tackle the problem of logical omniscience, as it allows, for instance, to distinguish between necessary propositions: while *1+1 = 2* and *Euler's Identity* are true in the same possible worlds (namely, all of them), intuitively what they are about, and how they are true at those worlds is different - the former seems to be partially about the numbers one and two, whereas the latter seems to be partially about pi and Euler's number. Here this approach won't be further developed and explored - that's the goal of the second part of the present dissertation. But even if one were to challenge the objection from compositionality's first assumption on the basis of a view of semantic content according to which what a sentence is about counts for its semantic value, that would not, however, help the case of the friend of impossible worlds. On most aboutness proposals, the Boolean connectives are transparent in regards to subject-matter, so that the topic or subject-matter of *p&q*, say, will just be the fusion of the topics of *p* and *q* taken in isolation, so that it would seem that the way *p&q* is true would also be a function of the ways both *p* and *q* are true, and therefore that in a world where the latter are true, both the truth-value of the conjunction and its topic should be a function of the corresponding value of its conjuncts. Aboutness approaches then seem to support a view according to which both components of semantic value should be determined by the values of its components in all worlds, contrary to what would be the case in impossible worlds in which conjuncts are true but not their conjunction.

Both this last conclusion, as well as the problem raised earlier for impossible worlds approaches rely, however, on the second assumption made explicit earlier: that the semantic values of non-modal complex sentences (that is, sentences formed from other sentences) should be determined by the values of those component sentences *in the same world*. This seems like a natural assumption, and in its favour it could also be said that since language users should be able to understand themselves, the semantic values of complex terms and sentences should be determinable from the values of their components at the actual world, for it seems to be their actual semantic value that is at stake when the respective values of the complex terms and sentences are under consideration. What would determine what "Water is H2O" actually means if not what the component terms actually mean? It would seem strange that to know what we actually mean when we utter sentences about water, we would have to first know what "water" means in a different possible world. It would seem, then, that if the friend of impossible worlds would like to reject the second assumption, they would have to do so in a way that still allows for the semantic values of complex terms and sentences in the actual world to be derived from their constituents. I believe such a way of rejecting the second assumption is available for a friend of impossible worlds, and that this path can be taken on the basis of some tentative points from Berto and Jago (2019) and Priest (forthcoming). In the next paragraph I try to present how this can be done. The question has nonetheless not been settled and more would have to be said about it.

Priest (forthcoming), in response to the compositionality objection to impossible worlds, states that what is needed for the learnability of a language is that it be recursive, that is, that it is possible to form more complex sentences and terms from simpler ones by the repeated application of certain semantic rules. And that, he goes on to say, is possible to do in an impossible worlds framework. I believe that Berto and Jago's (2019, pp. 180-184) reply to the present objection aligns with this remark by Priest, as it consists in showing how it is possible to form complex worldmaking sentences from the simpler ones they're made up from, and then arrive at the set of worlds of the former starting from the sets of worlds corresponding to the latter. Recall that the set of worlds that has been maintained to be a proposition is the content or semantic value of a natural language sentence as well as the content of agents' mental states. On the other hand, the lagadonian language sentences are not the content of the natural language sentences but provide rather the connections of properties and individuals of which worlds are sets and by which they can be identified. If, as it seems to be the case, we assume that more complex sentences of the lagadonian language can be formed from the simpler sentences in "roughly the way" the object language's sentences are[88], then it is possible to use this feature of the lagadonian language to get from the set of worlds containing certain lagadonian language sentences, to the set of worlds containing a sentence that binds the lagadonian language sentences with an operator (be it a logical connective, a quantifier, or a modal operator). Following Berto and Jago (2019, pp.182-183), this can be done in the following way. Consider an object language sentence, A, its lagadonian language translation A* and [A], the set of worlds at which A is true, where A is true at *w* if and only if A* ∈ *w*. If we take the intersection of the set [A], given how anarchic impossible worlds are, we simply get the singleton {A*}, so that "we can always move back-and-forth from the content (...) of an object language sentence A and the corresponding worldmaking sentence A*"[89]. Let us suppose, to simplify, that we wished to determine the set of worlds where ~A is true from the set of worlds where A is true. We can then take the intersection of [A], which is just {A*}, and arrive at {~*A*}, where ~* is just the lagadonian language term for the object language negation operator. This is done by concatenating the relevant symbols of the lagadonian language (which are themselves real entities, like sets, properties and individuals - for this reason, the concatenation is not actually performed, but is the mere possible combination of the relevant bits of reality). From {~*A*} we can finally get to [~A] by taking the union of {~*A*} with *w*, for every *w* in W (the set of all worlds), as since every set of sentences corresponds to a world, the union will just be the set of worlds at which ~A is true. We can, then, with the help of their lagadonian language translations, get from the semantic content of certain sentences to the semantic contents of the complex sentences formed out of them.

For Berto and Jago's proposal to work, the lagadonian language must be recursive, that is, it must be possible to "form"[90] complex sentences out of simpler sentences by the repeated application of certain rules, and, I take it, this is what Priest gets right in his short reply to the objection. Even if the sets of worlds at which sentences are true cannot be directly formed from one another if we consider impossible worlds, we can still get from one to the other indirectly if the lagadonian language is recursive. This, I take it, makes the proposal especially attractive, as even if, for independent reasons that have been considered earlier, the contents of sentences should not be considered to be structured entities, it seems that in order to understand what a given sentence means, that is, to know what set of worlds it corresponds to, it is necessary to have a grasp on what objects and relations are involved. Since worlds are just sets of sentences of the lagadonian language, then it seems that in order to know what set of worlds is a sentences' semantic value, that is, to know in what conditions it would be true, agents have to know to what

[88] Berto and Jago (2019, p. 183).

[89] *Ibidem*, p. 183.

[90] Here the quotation marks are used as it is not really the case that the sentences are formed, for the relevant objects and properties are not moved around, as it were, to form all the relevant sentences. Rather, these sentences all exist once we accept that the parts of reality out of which they are made of, exist and a principle of unrestricted recombination.

worlds belong certain connections of properties and individuals[91], for they are associated with the relevant terms of the object language sentence.

The approach can also meet the worry (whether it is legitimate or not) that the semantic values of complex non-modal sentences must be determined, in the actual world, by the semantic values of the simpler sentences in the actual world, in order for learnability to be maintained. It can do so in two ways. First, by stressing that the actual world is going to be one of the possible worlds, so that the actual world will always be a part of the semantic content of both the complex and simpler sentences whenever the value of the former depends on the values of the latter. Second, what is needed to get to the set of worlds on which a given complex sentence is true from the sets of worlds of the simple sentences that constitute it is, as we just saw, an understanding of how to "form" the more complex lagadonian sentence from the simpler ones, that is, to understand what sentences are concatenated with what operators and how they're so concatenated. However in order to do so one only needs to consider actual parts of reality to be differently combined. This is an aspect in which the lagadonian language's bridging of the gap between syntax and semantics is so helpful, in that the meaning of its terms are just the terms themselves, which are actual individuals, properties and set-theoretical constructions thereof.

It seems, then, that impossible worlds can be made sense of in a linguistic ersatzist framework and that the difficulties so far considered can be met. Besides some of the advantages already seen in accepting impossible worlds, the linguistic ersatzist conception of possible and impossible worlds might also have a big advantage in terms of theoretical economy. Impossible worlds are sometimes still perceived to be bizarre entities that we should not accept in our ontology, even if we already accept possible worlds. On the linguistic ersatzist view we can dispel these worries, both for impossible worlds and in regards to possible worlds. On this approach, worlds are no more than sets of sentences, sentences which are themselves nothing more than ordered tuples of actual individuals, properties, relations and set theoretical constructions. In a sense, one already accepts both possible and impossible worlds if one accepts in one's ontology particulars, universals[92] and sets, which are amongst the most basic ingredients in ontology. It immediately follows as well that if one accepts maximally consistent sets of sentences in one's ontology, then there is no reason to not also accept the existence of all other sets of sentences, so that there is no reason to accept possible worlds in one's ontology and forgo impossible worlds.

[91] More intuitively still, since the impossible worlds correspond to inconsistent and/or partial situations, this entails that knowing under what conditions is a given formula true, is to know in what situations the relevant connections of individuals, properties and relations are included.

[92] Here it is assumed that universals exist or that they must be accepted in order for linguistic ersatzism to get off the ground. That is not the case. That there are universals is just a simplifying assumption in the present context and the subjects so far explored are, for the most part, orthogonal to the millenial problem of universals. If the best solution for this latter problem turns out to be nominalist, then so the better for the linguistic ersatzist: instead of having particulars, universals and sets in their ontology, they can do without universals and only accept particulars and sets. If, further, one takes a reductionist stance on what sets are, as explored in the previous section, then even sets might be reduced to particulars.

Having extended the linguistic approach so far considered to accomodate for impossible worlds and having seen that no further difficulties arise, we can now move on to consider solutions that an impossible worlds' theorist might give to the problem of logical omniscience, and whether these are convincing. But before doing so, and in order to see how impossible worlds might help tackling the problem of logical omniscience, let us recall what the main difficulties were for the standard possible worlds analysis of intentional notions such as knowledge and belief.

All such worries stemmed from the fact that propositions were taken to be sets of possible worlds. For instance, on this approach there is *the* necessary proposition, which corresponds to all possible worlds, and *the* impossible proposition, which corresponds to the null set. Since there is only one necessary and one impossible proposition, it seems hard on this proposal to account for the fact that epistemic agents often have different attitudes towards what they take to be distinct necessary propositions, or distinct impossible propositions: for instance if someone knows that *2 + 2 = 4* but not *Goldbach's Conjecture*[93] (if we take the conjecture to be true), or a case where an agent does not believe that *there are round squares* but believes in *Goldbach's Conjecture* (if it turns out that the conjecture is false). More generally, the standard possible worlds framework has it that necessarily equivalent propositions are the same proposition, so that for instance *Batman is Batman* and *Batman is Bruce Wayne* turn out to be the same. This conclusion has (in other terms) been disputed by philosophers since Frege, for whom a proposition is a function of its constituent senses (*Sinne*), so that, given that "Bruce Wayne" and "Batman" have different senses, so do the aforementioned sentences, corresponding therefore to different propositions. The Fregean view assumes that proper names have senses, which is a very controversial thesis, but even if it is false, a defender of the view of propositions as sets of possible worlds should still explain how agents can have different attitudes towards one and the same proposition.

One way to do so is to refer to a theory of guises, where an agent might know that *p* under the guise of a given sentence, *s*, but not under the guise of a different sentence, *s'*. In cases like *Batman is Bruce Wayne* and *Batman is Batman*, the agents fail to have the same attitudes towards what is in fact the same proposition because it is presented in different guises. The appeal to guises can be considered as a part of a more broad metalinguistic approach: perhaps when an agent fails to know a given necessary proposition or a proposition necessarily equivalent to a proposition the agent already knows, what they actually fail to know is what proposition a given sentence expresses. As Stalnaker (1984) puts it, it is not the necessary proposition *p* itself that is the object of doubt, but rather the related proposition *sentence s expresses proposition p*.

This metalinguistic approach has something going for it. In the earlier case of *Batman is Batman* and *Batman is Bruce Wayne*, it might seem even to a friend of Russellian structured propositions that we are in the presence of just one proposition, as intuitively the individuals and

[93] That is, the famous conjecture in Number Theory that states that all even numbers greater than two can be expressed as the sum of two prime numbers.

relations involved are the same[94]. It might be said that in such cases the agents fail to know what "Batman" (or "Bruce Wayne") refers to, so that they do not know what proposition is expressed by "Batman is Bruce Wayne" and that if they knew what proposition it was, the statement would be just as informative as the proposition expressed by "Batman is Batman".

However, it is doubtful that this solution works for all the cases where agents seemingly fail to know a given necessary proposition. As Jago (2014) claims, it seems that some highly non-trivial theorems of Mathematics can be stated in terms that can be easily understood by agents who have had some formal education in the discipline. For instance, Fermat's Last Theorem can be stated using notation that a good number of high school students are already familiar with: there is no natural number n greater than 2 and integers x, y and z for which the equation x^n + y^n = z^n has a solution. It seems, moreover, that the theorem can be stated in a very simple way, in just a couple of lines, so that there seems to be no good motivation to count this among the cases Stalnaker presents where high structural complexity and the use of nested logical connectives make it plausible that agents fail to know what proposition is being expressed by a given sentence, even if they're familiar with the notation employed. Similarly, mathematicians seemingly were not in the same situation as someone who fails to know what "Batman is Bruce Wayne" expresses because of a failure to know what some of the terms refer to. Even though what numbers and mathematical objects and structures are is still a matter of great philosophical discussion, it does not seem plausible to claim that the reason mathematicians struggled to prove Fermat's Last Theorem for three centuries is simply because they did not know what the natural numbers, the integers or powers' nature is. If that were the case, and given there is still great controversy over these questions in the Philosophy of Mathematics, it is not clear why mathematicians would now be in a better position to know what is expressed by a statement of Fermat's Last Theorem than they were before it was first proved.

If these considerations are on the right track, then it seems that a metalinguistic approach cannot be applied successfully to all instances where agents fail to know necessary propositions, or one of two necessarily equivalent propositions, while knowing the other. If all agents already knew the proposition expressed by Fermat's Last Theorem, and if the theorem can be stated in a simple and unambiguous way, it is therefore not clear why it was treated as a mere conjecture by mathematicians until a proof of it was first presented, and why such a proof was so sought after. Provided no further reasons are given to consider this, and other such cases, as arising from mere ignorance about what sentences express what propositions, then it seems that we should accept the intuitive explanation that agents did not in fact know Fermat's Last Theorem for why they sought a proof for it, and we should therefore aim to provide a framework on which agents can

[94] This specific case is more complicated than the current description, both because Bruce Wayne, and Batman, are fictional characters, and because in the fictional universe to which they belong different characters have played the role of the Batman (for instance, Dick Grayson, who is more well-known as the original Robin in the comics). Here the example is used simply because of how famous the fictional character is, as well as the fact that his identity is kept hidden to most other characters in the fictional universe. The complicating factors just mentioned are here being ignored. If the reader is unhappy with the example, they're free to consider instead *Lewis is Lewis* and *Lewis is Bruce Le Catt* as simpler, and actual, examples of, respectively, uninformative and informative identity statements containing only proper names.

come out as failing to know necessary truths. In order to do so, we should go hyperintensional, and a natural way of doing so is to just expand the possible worlds framework with impossible worlds, changing it as little as possible.

The impossible worlds approach to logical omniscience

Having considered how to expand the linguistic ersatzist possible worlds' framework to accommodate impossible worlds, and replied to some general objections against this strategy, we now move on to consider how to develop in detail the impossible worlds' solution to the problem of logical omniscience, what objections it faces and how these can be met, or not.

If the intuitions about cases such as that of an agent who wonders whether Goldbach's Conjecture is true are on the right track, then some impossibilities are epistemically possible for some epistemic agents. These impossibilities in turn correspond, on the present approach, to impossible worlds, which are themselves simply inconsistent and/or incomplete sets of sentences of a lagadonian language. However, not all impossibilities are born the same, and while some seem to be epistemically possible for some agents, others are not epistemically possible for those same agents, and, perhaps, for any epistemic agents whatsoever. For instance, it might not be epistemically possible for any agent that there is a geometrical figure that is at the same time a square and round. The epistemic possible worlds cannot, then, be identified with the whole realm of worlds that we have so far accepted into our ontology (albeit in the ontologically innocent way that they reduce to set-theoretic constructions of recombinations of actual bits of reality). The first and primary task for a defender of an impossible worlds approach is, then, to delimit the class of all worlds to what we might call "epistemic space" (Chalmers (2010)).

In order to do so, let us first consider if we should restrict the realm of worlds for what is epistemically possible unrestrictedly, that is, what is epistemically possible for any epistemic agent whatsoever. Jago (2014) considers that blatantly inconsistent worlds, that is, worlds where both a proposition and its negation are true, should be ruled out as epistemically possible for any agents whatsoever[95]. This seems to go in line with the case of the round square, where the property of being square, by its definition, entails that a square is never a round geometrical figure. Lewis (2004) in the same vein claims that Priest's dialetheist tools allow us to reason about what's blatantly impossible, whereas what seems to be the case is that we reason rather about what is simply subtly impossible, that is, cases which, in spite of being impossible, we fail, when we reason about them, to see that they are so. I take this position to be very intuitive and to correspond to what in fact goes on in human agents' epistemic lifes. It also seems to be the case that the same should apply to the case of even more limited agents. In fact, if for instance we take the normativity of logic to be constitutive of thinking - that is, if we take thinking to be an activity that is defined by being answerable to the principles of logic (MacFarlane (2002, p. 37))-, then it seems that these norms will preclude systems who constantly consider blatantly impossible scenarios, that is, that ignore the law of non-contradiction, to be considered to be engaging in thinking.

[95] Berto and Jago (2019) add to the restriction that worlds representing the respective conjunction should also be ruled out from epistemic space, so that worlds representing A & ~A are directly ruled out from epistemic space. But even if this restriction is not added, as will soon become clear, there is a way to maintain that worlds representing A & ~A are easily ruled out from epistemic space for minimally rational agents. Still, this restriction would have it that worlds representing A & A and ~A would not be ruled out from epistemic space.

Here an objection is that it seems that dialetheists, such as Graham Priest, may be misguided in their beliefs, but they are not, still, irrational, or at the very least to the point where they might be said to no longer be engaged in the activity of thinking. I take Jago's (2014, pp. 221-226) reply to be perfectly satisfactory at this point: the reason why dialetheists may believe in contradictions and yet contradictions are blatantly impossible and therefore not epistemically possible is that dialetheists do not interpret negation as it is classically understood and therefore do not believe contradictions in the sense just considered. A cursory read of the literature on negation, and the debate between dialetheists and friends of classical logic reveals that to a great extent they do not agree on how the negation connective is to be interpreted. Here when we maintain, then, that a blatantly inconsistent world is not epistemically possible because it contains a contradiction, that is so if we take negation to be classical, whether or not it correctly aligns with the meaning of "not" in everyday English. When dialetheists claim to believe contradictions, we should then take them to believe contradictions*, where to believe a contradiction* is to believe that A & ~*A, instead of A & ~A, where ~* corresponds to the dialetheist's interpretation of the negation connective. We can then proceed by taking worlds in which contradictions are true to be blatantly inconsistent and therefore to be epistemically impossible.

Jago (2014) and Berto and Jago (2019) then propose to limit the epistemically possible worlds by linking impossible worlds to blatantly inconsistent worlds. Let us consider as an example the following world, w = {A, A → B, ~B}. This world is not itself blatantly inconsistent, as it does not contain any lagadonian sentence and its negation. It is, however, very close to being one, as we can get from w to a different world, w^* = {A, A → B, B, ~B} by one application of the *modus ponens* rule of inference, where w^* is a blatantly inconsistent world. In general, for logically competent agents such as the average human being, worlds such as w won't be epistemically possible as they can easily dig out, as it were, the "hidden" contradictions that they contain. For this reason, Berto and Jago then go on to define epistemic possibility in a way that precludes worlds such as w, that is, worlds from which, by a small number of applications of certain rules of inference, from counting as epistemically possible. Since what counts as a "small" number of steps is vague, Berto and Jago take epistemic space itself to have vague boundaries - it is not determinate what is epistemically possible. I believe that we do not have to make this move at this point. Perhaps for agents such as the average adult human being, worlds that contain hidden contradictions that can be easily dug out are not epistemically possible, but perhaps for less logically competent agents, such worlds turn out to be epistemically possible. I will assume, then, that epistemic space has determinate boundaries: all worlds are in, save for the blatantly inconsistent ones, which are the worlds representing contradictions, either by representing both of a contradiction's conjuncts, or by representing the conjunction itself.

Berto and Jago present an ingenious way of modelling the process of digging out hidden contradictions from subtly inconsistent worlds. They start by taking an inference rule, such as modus ponens where A, A → B ⊢ B, and move from a world in which the premises are true but not the conclusion to the union of that world with a world representing the conclusion. If what a world represents is inconsistent, then by this procedure we will at some point arrive at a world that contains both a lagadonian sentence and its negation, that is, a blatantly inconsistent world. We can then define the rank of a world relative to a given set of inference rules to be the minimal number of steps, using those rules, required to arrive at a blatantly inconsistent world. The idea is then that the higher the rank of a world, the harder it is to realize that it is inconsistent, with blatantly inconsistent worlds having rank 0 and possible and incomplete but consistent worlds having an infinite rank, so that they're deeply epistemically possible, that is, they are epistemically possible for any agent, if we don't take into account their informational states[96].

With the notion of a rank in place, Berto and Jago go on to define deep epistemic possibility in terms of possible worlds, incomplete worlds and subtly inconsistent impossible worlds. What counts as a subtle impossibility seems, as Lewis stressed, to be context-dependent and vague. It seems, then, that we can't define the subtly inconsistent worlds as the worlds of rank higher than *n* for a given specific *n*. For this reason, Berto and Jago take the approach that what worlds are subtly inconsistent, and therefore the extension of the epistemically possible worlds, is vague[97], as there is no value for *n* such that it is determinate that agents rule out all worlds of rank lower than *n*. This vagueness in turn also allows Jago (2014) to provide a solution for what he calls "Bjerring's Problem", that can be posed as a dilemma: either we only accept possible worlds, case in which we have to characterize agents as being logically omniscient; or we add impossible worlds, case in which we can describe agents that are not logically omniscient, but that are wholly logically incompetent; one way or the other, it is not possible to characterize, in this framework, limited but logically competent agents, which was the goal from the get go. So far a lot has already been said about the first horn of the dilemma, so let us focus on the second horn.

As Bjerring (2012) argues, any minimally logical competent agent should believe the trivial consequences of what they believe. For instance, it seems that any minimally logically competent agent should believe a conjunction whenever they explicitly believe the conjuncts, or the consequent of a conditional when they explicitly believe the antecedent and the conditional, if we take these to be trivial inference rules. It seems, then, that the agents' beliefs should be closed under trivial logical consequence. But even the most complex inferences can be expressed as sequences of trivial inferences, so that if belief is closed under trivial logical consequence, as it seems to be required by logical competence, then belief is closed under logical consequence *simpliciter*, so that we either have logical omniscience, or no logical competence at all[98]. Jago's

[96] In Chalmers's (2010) description of the notion, it corresponds to what can't be ruled out by *a priori* reasoning alone. I take both characterizations to be equivalent.

[97] Here note that what has been so far characterized is what it is for a world to be blatantly inconsistent, and therefore blatantly impossible, as opposed to subtly impossible, for all agents. Depending on how agents reason, their informational state and their cognitive capacities, then they'll be able to limit epistemic space even more so what comes out as epistemically possible for a given agent will be a subset of the worlds in deep epistemic space. It is this notion of what is epistemically possible for agents that Berto and Jago consider vague, given that it is also vague from what "epistemic oversights" agents suffer, so that they fail to draw all the logical conclusions of what they know.

[98] Here note that the claim is not that chains of trivial inferences are themselves trivial, but rather that all complex inferences can be presented as such chains of trivial inferences. Given that all steps in the chain are trivial, Jago

(2014) ingenious solution for the present difficulty has it that it is never determinately true what trivial logical inferences logically competent agents fail to draw, that is, what specific *epistemic oversights* they suffer from, even though it is the case that, if they are not logically omniscient, they must fail to do so at least for one trivial logical inference. In this way, vagueness is also part of the picture when it comes to what inferences are trivial and what inferences are not, so that in some cases it will also be indeterminate whether a given impossible world is epistemically possible or not for a given agent.

Incomplete worlds

As we just saw, in Berto and Jago's view, incomplete but consistent worlds are deeply epistemically possible. As they state (Berto and Jago (2019, p. 225)), for instance we do not consider the Sherlock Holmes stories to be impossible simply on the basis that in the stories it might be true that the detective had breakfast on a particular day, though nothing is said about what the breakfast consisted of. They take this to show that when considered as epistemic possibilities, incomplete worlds should be taken as incomplete representations of complete states of affairs, as long as they are not inconsistent. This move, however, seems to be contrary to the option taken earlier to consider incomplete but consistent worlds to be impossible worlds, for if they are taken to just be representations of complete worlds, then they are not being considered by themselves in our characterization of epistemic possibility. Of course, Berto and Jago might claim that the problem of logical omniscience and how to characterize epistemic space is far from being the only application that there is for impossible worlds, including incomplete but consistent worlds, and that it is only in the particular case of epistemic possibility that incomplete but consistent worlds do not play a role by themselves, but only as representations of complete worlds. The argument they provide for this view, however, seems lacking, in that: i) it does not seem like the case of stories generalizes for the case of worlds, for it is a highly contested matter what stories represent implicitly and how they relate to common belief, common belief which, presumably, includes the belief that a full description of reality is not incomplete; and relatedly ii) it seems arbitrary to take incomplete worlds to have different representation conditions, instead of behaving as all other worlds, which represent that something is the case by containing the appropriate lagadonian sentences as members. If this is right, then stories, but not worlds, represent implicitly, and therein comes the disanalogy and the reason why we do not consider the Sherlock stories to be impossible simply on the count of nothing being said about every last detail of its world.

In fact, I believe that if we delve a bit deeper and explore more carefully the case presented to motivate the idea that incomplete but consistent worlds are always deep epistemic possibilities, we will find that, contrary to what Berto and Jago argue, they provide a case *for* considering incomplete but consistent worlds to, in some cases, not be epistemically possible. To see how this could be, consider the question of *why* it is that, when confronted with incomplete

claims, then it is never determinate at what step in the chain did the agent fail to reach a trivial conclusion of some premises they started from. The claim therefore is the exact opposite: that chains of trivial inferences are not themselves trivial for limited agents, and that the difficulty comes exactly from this conflict between the triviality of the steps and the highly non-trivial nature of some of the inferences they add up to.

worlds, agents tend to, as it were, fill in the gaps. A plausible hypothesis, I take it, is that just like when given contradictory information, agents try to find strategies to integrate the information they have been given in a way that avoids contradiction - for instance if one gets information about the round square, a principle of charity might require the agent to reinterpret the expression "round square" so as to make sense of what the speaker is saying -, perhaps agents fill in the details of incomplete information they are given about alternative scenarios (or consider various ways in which they might do so) exactly because they know that the world could *not* have been incomplete in that way, that something else besides what is given must be the case. In both cases, the process can be so automatic that the agents might even fail to realize what strategies they are deploying in order to successfully integrate information.

Let us consider two examples, one for inconsistency and one for incompleteness. Suppose that Rita, José and Mateus are going to have dinner, but that Rita does not know yet whether Carla will join them or not. Upon asking one of her friends whether Carla will join them for dinner, she gets the answer "Yes, and no.", using the information she has available regarding the plans for the night, Rita concludes that her friend must mean that Carla is coming with the group, but not having dinner with them[99]. Now suppose that everything is as in the last case, but upon questioning one of her friends, she gets instead the reply "She's on her way, but her mother clocked off early from work today." Upon hearing this, and in order to make sense of the seemingly irrelevant information she has just been given, Rita concludes that Carla must have had dinner at home, but is joining them anyway. In the first case, Rita gets seemingly contradictory information about whether Carla is joining the group for dinner, and must therefore, to make sense of the information being given, reinterpret the speaker in a way that their answer is not a blatant contradiction, but rather a statement where the "Yes" and "no" stand for the affirmation and negation of different propositions. As for the second, it seems that Rita must fill in the details in order to make the information she has just received relevant to the goal of the conversation - this filling in the details, I take it, involves a complex exercise of deploying relevant background information, such as that Carla's mother has cooked dinner for her, that Carla does not have dinner twice, that by clocking off early Carla's mother will have time to cook dinner, and so on, including general information about the world and its composition. It seems that only by adding all these details it is possible to find the information provided relevant for the goals of the conversation.

It would seem, then, that this act of filling in the details is a very common practice for human agents, one that is probably automatic in a number of contexts. It is because of how natural this act of filling in the details can be that, I take it, Berto and Jago were inclined to maintain that incomplete but consistent worlds are not epistemically impossible. But if anything, I believe examples like this show, rather, that human agents might be so used to considering cases like this impossible and that something else besides what's given *must* be true that they

[99] Here there would be plenty of ways to make the case more precise, in regards to what information Rita has available. For instance, it could be that there were only three options: Carla goes and has dinner with the group; Carla goes but does not have dinner with the group; Carla doesn't go. Case in which it is true *in a sense* (as it is usually said) that Carla is *joining them* for dinner, but not true that, *in a different sense*, she's joining them *for dinner*.

even fail to consider the cases of blatantly incomplete worlds by themselves, without adding information to them from the common ground, information that the relevant worlds do not contain. Once one focuses on just what sentences a world contains, putting aside the natural tendency to add information that the world does not contain, then, I take it, such worlds are automatically ruled out by agents as impossible, when considered as candidates for full descriptions of actuality. This ruling out of incomplete worlds as epistemically impossible explains why we tend to so rarely reason about such worlds without adding any information to them[100].

If the preceding considerations are right, then besides blatantly inconsistent worlds, we should also rule out blatantly incomplete worlds from epistemic space. So for instance worlds in which it is true I had breakfast but there is no fact of the matter as to what in specific I had for breakfast, are worlds that agents automatically rule out as impossible as soon as they consider that given that I had breakfast, then it *must* be the case that I had something or other for breakfast[101]. This, however, introduces a much greater degree of complexity to the model, as again it won't always be clear whether a world is blatantly or only subtly incomplete and how that can be ascertained. Here I believe a full answer would necessarily have to start off from the best evidence available from cognitive psychology. Still, we might consider several factors that might be at play. For instance, perhaps agents more easily notice incompleteness if the lagadonian language sentences missing from a given world have the same subject matter as some

[100] Here an objection to this way of motivating why blatantly incomplete worlds are ruled out from epistemic space is that if correct, then the argument just provided would overgenerate, as perfectly reasonable answers to the question of whether Carla is coming over for dinner could then not be believed by agents, as they would have to start adding details such as what clothes Carla would be wearing, if she would be wearing glasses, and so on. This objection highlights a disanalogy between the example just considered in which some background information must be added so as to make a certain contribution to a conversation relevant and the case of incomplete worlds. Anticipating some of the topics that will be discussed more thoroughly in the second part of the dissertation, we might say that when considering the question of whether Carla is coming for dinner, the agent wants to know whether the actual world is among the worlds in which Carla joins for dinner, or otherwise among the worlds in which Carla does not join for dinner. The agent therefore is not interested, in respect to the relevant question, in knowing what the actual world is in full detail, but rather in what group (cell of the partition, as we will see later) it is situated. Perhaps a better analogy at this point would be if an agent got an incomplete answer to a given question and were then told that there was no fact of the matter as to what complete answer is the correct one. So for instance if an agent wants to know what Carla's profession is, and is told that Carla is either a teacher or a software engineer, it seems that they thereby restrict their attention to worlds where Carla is a teacher and worlds where Carla is a software engineer. And here it seems that worlds representing simply one of them but not which are excluded as agents know (let us suppose) that for Carla to be either a software engineer or a teacher, she must thereby have one profession or the other.

[101] Here the incomplete worlds are automatically ruled out, it is claimed, because they are obviously incomplete as candidates for complete alternatives to actuality. But, it might be claimed, for this to be the case, perhaps we are tacitly assuming that worlds include a totality sentence in the lagadonian language expressing that nothing else besides what the world represents is the case. And if such a sentence were being tacitly included, then perhaps the worlds would then cease to be ruled out because they were incomplete, but rather because they are inconsistent. I believe that no such sentence has been tacitly added. By characterizing a given situation as a world we are *ipso facto* considering what it represents as a complete alternative to actuality, whether possible or impossible, and that is so regardless of whether the situation also contains a totality fact or not. This aligns with the intuitive understanding of worlds as ways for the total universe to be, and the current proposal is simply that besides being inconsistent, some situations are ways that the world could not have been simply because the world is complete and some such situations are partial.

of the sentences that are members of that world; or when the number of sentences contained in a world is small; or when the missing sentences would describe a spatial region contiguous to that being described by some of the other sentences in the world; whether an agent's awareness has been brought to bear on the missing sentences, among many other such factors.

Despite a world's being ruled out as blatantly incomplete by an agent likely depends on a complex interplay of various factors, it might still be, however, that in all such instances a general description applies - that is, that regardless of how an agent came to take notice of the fact that the world is incomplete, that the mechanism by which they rule them out as epistemically impossible is still, at a given level of generality, the same. And, given what has been said earlier, I believe one might take an agent's filling in the details or adding information that a world does not contain, to be a clear indicator that a given world has been ruled out on the count of being incomplete, for when they do so and take that some such information *must* be added for the alternative scenario to be a full description of reality, then they have ceased to consider it to be a genuine alternative for what reality, in its entirety, may be like.

This description, however, only tells us when a given agent rules out an incomplete world as epistemically impossible, not when such worlds are blatantly epistemically impossible. Here, however, I think the situation is similar to the case just seen of blatantly inconsistent worlds. We can consider the empty set of sentences to be the most blatantly incomplete world, so that for no agent is the empty world a candidate for a way the world may be for all they know, which is tantamount to the view that for any agent, there not being any true proposition is ruled out as a way that the actual world may be, as for instance they might at the very least know that they exist[102]. We then need to find a way of linking the incomplete worlds to the blatantly incomplete worlds, similarly to how Berto and Jago have linked inconsistent worlds to worlds explicitly representing contradictions. I propose that we do this by considering the pair of intersections of a given world with the intersection of all worlds in *p* and the intersection of all worlds in *~p*, for all propositions *p*. If for any such pair, both intersections are the empty set, that is, the most blatantly incomplete world, then the world is incomplete. If the result in non-empty, then we will arrive at the singleton of the lagadonian language sentence corresponding to *p* or to *~p.* We can then define the rank of completeness of a world to be equal to the number of propositions where it is the case that one of the members of the pair of intersections mentioned above is not empty. So for instance the empty set will have rank 0, for there is no proposition that, either the intersection of worlds in it or its negation, intersected with the empty set will result in anything other than the empty set, the sets {A} and {~A} are both worlds of rank 1, for there is only one

[102] This is, admittedly, a Cartesian view. Taking the empty set of sentences to be blatantly incomplete and therefore to not be deeply epistemically possible does not, however, depend on this Cartesian thesis. One might for instance think that all epistemic agents should know the law of identity, so that only words containing the lagadonian sentence corresponding to *For all x, x=x* should count as epistemically possible. The view that the empty world is deeply epistemically impossible only requires that all epistemic agents have at least one item of *a priori* knowledge.

proposition for which there is an intersection that does not result in the empty set[103], and so on, until presumably complete worlds, which will have an infinite rank.

Now armed with a parallel notion of rank for incompleteness, we may then say that some worlds are not deeply epistemically possible when they are either blatantly incomplete or blatantly inconsistent, as well as that worlds may fail to be epistemically possible for a given agent both because they're incomplete and because they're inconsistent. In principle we could also apply the vagueness solution for the case of completeness of worlds. It might seem, though, that intuition here isn't as helpful, for it might not be clear that for a given incomplete world, say a world that only contains three lagadonian language sentences, the agent should easily see that one further sentence is missing. If this is so, then we would not be ascribing irrationality to the agent if we described them as failing to realize that such a world is incomplete. Still, perhaps when worlds get closer to being complete, it gets harder for agents to know whether a further sentence is missing. We could, then, perhaps, in principle, follow the same strategy that Berto and Jago have followed and claim that these correspond to oversights on the agent's part, so that it is not determinate at what point the agent has failed to realize that a given world is incomplete.

Here a number of issues arise. One might think that taking the rank of completeness of a world to just be given in terms of the number of propositions that the world belongs to (equivalently, to be given in terms of the number of lagadonian sentences a given world contains) is an oversimplification, since even if worlds with a larger number of lagadonian language sentences might tend to more easily give off the impression that they're complete, other factors, such as the ones mentioned above, might also be relevant in determining whether a given incomplete world is higher or lower in the completeness rank, that is, they might play a role in determining how easy they can be spotted by agents to be incomplete. I believe such worries are totally fair, and what I have to say at this time about it is not entirely satisfactory.

The first point I'd like to make could be taken to be a *tu quoque*, were it not for the second point, and it consists in stating that Berto and Jago's approach likewise simplifies matters and has some level of idealization built into it, namely in regards to what rules of inference agents deploy in order to uncover hidden contradictions in worlds, and by considering that only the number of steps matter. For instance, is it clear that a world that can be linked to a blatantly inconsistent world by seven applications of *modus ponens* is less likely to be ruled out by agents as a candidate for actuality on the count of being inconsistent than a world where it takes 6 steps to get to a blatantly inconsistent world, but the relevant sequence of steps involves a very specific combination of several different inference rules? My intuition is that this is not the case, or at least not in general the case. Berto and Jago's answer also takes all primitive inference rules to have the same computational costs, merely being sensitive to the number of applications of the rules, while this might not be the case. For instance if we allow disjunction elimination to count as one of the rules applied by agents, then perhaps it'll have a greater computational cost than rules such as conjunction elimination, yet Berto and Jago's model takes all such inference rules

[103] Here notice that the blatantly inconsistent world {A, ~A} is also a world with a completeness rank of 1, while having a consistency rank of 0. In general the values of the rank of worlds for completeness and consistency will be independent of one another. Possible worlds have the maximum value in both ranks.

to be the same when it comes to the matter of what worlds count as epistemically (im)possible. Likewise, to suppose that agents reason in accordance with the rules of a given natural deduction system seems to correspond to an idealization. Perhaps, even restricting ourselves to human agents, it is not determinate what rules of inference a given agent utilizes, and these might be volatile and very dependent on heuristics (Williamson (2020) explores some of the ways in which heuristics might play an important role in reasoning).

The second point is that in both cases, these do not seem like devastating objections that show that the entire project is misguided or based on unsound footing, but merely that more work needs to be done to explore all these intricacies and provide a more realistic picture of how it is that agents rule out worlds as epistemically impossible, something that can only be done, I take it, on the basis of a careful consideration of work done in cognitive psychology. Here I won't try to provide such a fuller and more realistic account of how to characterize the notions of a blatantly or subtly incomplete world, as that would be a subject worth dedicating an entirely different dissertation to. Instead I'll focus on the impossible worlds approach as it has been so far explored and presented, considering objections raised for it even in these admittedly simple and idealized versions.

It was just seen that appealing to vagueness in the case of incomplete worlds does not seem nearly as intuitive as in the case of the successive application of trivial inference rules. To see why, consider that it does not seem in general to be the case that if an agent fails to notice that neither of the members of a pair formed by a proposition and its negation is represented by a given world, that is some form of irrationality on their part. For instance, it might be possible that there exist elusive objects, that is, objects which, given their dimension and properties, are in principle undetectable for human-like agents with the perceptual mechanisms and tools we possess. Suppose then that a given human agent is considering two worlds, which are wholly the same, save for the fact that one does not represent anything about particular elusive objects, whereas the other includes lagadonian sentences expressing what properties and relations each elusive object instantiates. Given that the agent has no evidence whatsoever in regards to the existence of these objects and what their particular properties are, then she might consider the first world to be a complete description of reality, as her evidential background does not allow her to distinguish between the two worlds. It seems, then, that to attribute to this agent the failure to see that the former is incomplete is not to attribute to the agent any form of irrationality - in fact, the opposite might be true, as it seems that there would be in principle no reason for the agent to consider the former world to be incomplete, even though it is in fact so. It would seem, then, that contrary to the case of inconsistent worlds, taking the case of what worlds an agent is able to figure out are incomplete to be a vague matter such that no specific oversight can be attributed to agents on the pain of ascribing irrationality to them is much less compelling. However, it seemed like the failure of ascribing to agents specific oversights was one of the major motivations for appealing to a solution based on vagueness.

Going dynamic

Giving up on accepting a vagueness approach for the case of incompleteness seemingly has two drawbacks. The first one is that we then cannot easily explain why it is hard to say where determinately agents fail to consider worlds as incomplete, that is, of why it does not seem to be feasible to draw a sharp line between what worlds an agent is somehow led to think are complete, and those that they realize are indeed incomplete. The second is that we lose what has been so far a symmetry between the cases for inconsistency and incompleteness.

I believe that there is a single reply to both drawbacks: that both in the case of inconsistency and in the case of incompleteness, we should not go for an approach based on vagueness, but rather for an approach based on dynamic epistemic logic. I try to motivate this approach in the following paragraphs, starting by considering the case of inconsistency and the problems I think can be presented to Jago's move.

Bjerring and Skipper (2018) criticize Jago's (2014) solution from vagueness to what he called Bjerring's problem, that is, the dilemma that it seems that in an impossible worlds framework it is not possible to avoid the two unattractive extremes that human agents are wholly logically incompetent, for they fail to know what trivially follows from what they know, and that human agents are logically omniscient, knowing everything that follows from what they know. Part of their critique amounts to the claim that attributing an epistemic oversight to an agent does not amount, thereby, to treating that agent as irrational, for even if they fail to derive a trivial consequence of what they know, they might still have the ability to do so, and it is the possession of this ability that makes agents logically competent and minimally rational without being logically omniscient.

An immediate issue that arises from these considerations is how to describe this ability, so as to then say what constitutes an agent's being minimally rational. Bjerring and Skipper (2018) envision a behavioural test: if one knows that *p,* and *q* follows trivially from *p* by one application of given rules, upon being asked whether *q* is the case, does the agent immediately answer that yes? If the answer is positive, the agent passes the test. On the other hand if the answer is negative, then the agent fails the test and is not logically competent. It is not clear whether Bjerring and Skipper's test is convincing as an infallible test. For instance, if the test-taker knows that *p* but knowledge, as Williamson (2000) has argued, is not luminous, they might fail to know that they know *p,* or even fail to be aware of it. In the case where the agent is asked whether *q* is true or not, if their knowledge of one of the premises from which *q* is derived is not transparent to them, then perhaps the agent will not assent to *q*, even though *q* trivially follows from what they know. Of course, Bjerring and Skipper might take this to just be one of the cases where we show that an agent is not logically competent. But if that's so, then logical competence becomes impossible to achieve, at least for the average human agent, and it was exactly such a minimal, achievable, and often achieved, level of logical competence that we set out to characterize in the first place. It would seem, then, that Bjerring and Skipper's test cannot serve its purpose.

Besides leading to the wrong result when agents are not aware of what some of their items of knowledge are, it would seem that likewise a myriad of other situations could arise that would lead logically competent subjects to fail the test. Bjerring and Skipper (2018) mention in an indirect way some such situations, when giving an example of an application of their test they say: "Assuming that you are attentive, mentally well-functioning, and so on" in regards to whether a rational agent would reply in the requisite way. Returning to a previous point about metalinguistic doubt, the test-taker might fail to realize that a certain sentence they have been presented with refers to *q*, so that they interpret the question in an unintended way, so that the question does not prime them to derive *q* from *p*. Similarly, even if the test-taker's attention is brought to bear on the question of whether *q* successfully, they might not be, for a variety of reasons, focused on the test to the extent that inferring *q* from *p* requires. Perhaps some other factors not yet mentioned, such as the test-taker not being sincere, or what logic they are operating with would also play a role in determining whether a logically competent agent would pass the test or not, but in any case, it seems plausible that at least in some rare circumstances, they will fail to pass the test.

Despite the difficulties just identified for taking Bjerring and Skipper's test to play the role they wish for it to play, perhaps it is still the case that the test can help us establish whether a given agent is logically competent or not. Two ways of avoiding these difficulties run as follows. The first is to consider not how agents actually reply when put to the test, but rather how they would reply in optimal conditions, which presumably include the fact that their cognitive systems are working well, that they're fully focused on the test, that they understand what the question being asked amounts to, that they're truthful in their answers, and so on. The second way is to consider what agents actually do, but instead of taking an agent to be minimally logically competent if they answer in the expected way in all instances, to take them to be so competent if they reliably act in the expected way. A logically incompetent agent would perhaps hit and miss roughly half of the time, given they could not know whether *q* is true or not, and would take a guess, say, whereas logically competent agents, despite the fact that they might fail the test on some occasions for some of the reasons so far considered, will tend to reply in the appropriate way. I believe that either of these options are plausible, and so that we might take a strategy in the vicinity of what Bjerring and Skipper propose to characterize the notion of minimal logical competence.

With the revised way of understanding the role of the behavioural test proposed by Bjerring and Skipper, we are now in a position to criticize the motivation for treating what counts as epistemically possible for a given agent to be inherently vague. When an agent fails to draw out some conclusions from what they know, attributing to them such a failure does not amount to treating that agent as irrational, but rather to say that they, for some reason or other, have failed to perform in accordance to one of their abilities, logical competence. Even if it is an agent's logical competence that allows us to attribute to them minimal rationality, if logical competence is understood as an ability, then pontual failures to live up to its standards do not preclude agents from being considered minimally rational. Perhaps in this aspect rationality is not so different from other notions that can exert some form of normative force on agents, such as fairness or kindness, to the point where such attributes are valued and represent what individuals should strive to attain. So for instance, just as a kind person may not act kindly in all circumstances, or a fair person may not act fairly towards others in every context, perhaps a minimally rational agent may fail to display logical competence on certain occasions without thereby ceasing to be rational. Of course, kindness and fairness are, perhaps like rationality and logical competence, gradative notions, as agents might be more or less fair or kind, and one way for agents to so is to act kindly or fairly more or less often, and in this sense we can then distinguish between the ideal of always meeting the standards of minimal rationality and the extent to which human agents do so and thereby count as logically competent. Here the point is not that there are sharp boundaries in regard to whether an agent counts as rational or not, but rather that perhaps the notion works as other gradative notions such as fairness and kindness in that attributing a failure to live up to certain standards on certain occasions, which in the present context of logical competence means attributing a failure to derive a trivial consequence of given premises, need not entail that an agent is thereby irrational or is not logically competent.While it might be the case that it’s very difficult, or even practically impossible to ascertain exactly what epistemic oversights a given agent might suffer from on any given occasion where they fail to draw all the conclusions from information they have available, still these difficulties do not show that what counts as epistemically possible should be itself a vague matter.

An important aspect about this last practical impossibility has to do with what may be the most natural way of trying to ascertain where an agent might fail to draw out a trivial consequence of what they know: to ask them questions about what they know, and from there derive what was the point in the chain of steps they took to derive a given conclusion from known premises where they failed to take the next one. Bjerring and Skipper’s remark suggests that they hold that asking questions prompts agents to reason in certain ways, and I believe they’re right in doing so, and similarly that philosophers like Yalcin (2018) and Hoek (forthcoming) are right in holding that the notion of belief is closely tied to that of a question. Agents are not just limited in terms of their ability to compute the consequences of what they know, but also in how much information they can store and how that information is accessed - as we previously considered, it might even be that information, for limited agents, is stored in different fragments, with what fragment a given piece of information gets retrieved from depending on what information the agent needs at any given time. Asking agents questions about the contents at issue will bring their awareness to them, and that will prompt them to retrieve information associated with the subject. A classical example of how questions can guide inquiry is present in Plato’s Meno, where in the process of trying to prove to his interlocutor that the knowledge of the Forms already exists in the mind before an agent gets educated, Socrates (the character) queries a slave about geometry, and gets him to prove a theorem that the slave had never deduced.

That inquiry is guided by what questions agents are led to consider is one of the ingredients of what will come to be the proposed way of dealing with the problem of logical omniscience, and will be further explored in the second part of the present dissertation. For now, however, it is important to stress how the picture just roughly sketched relies heavily on a dynamic understanding of what epistemic agents come to know. If the preceding considerations are correct, what questions agents are confronted with creates a new context, a context in which they join information they may even have stored in different fragments, and in which different consequences of what the agent knows become more salient, and therefore more easily deducible. It is against the background of this dynamic process that epistemic possibility should be considered. In fact, that the process is so dynamic might be one of the reasons behind why it seems intuitive to a certain extent that what is epistemically possible for a given agent in any given context is vague, for it will usually be impossible to have the agent in the exact same context they once were, and to, as it were, crystalize that context to then ascertain at what point the agent failed to derive certain consequences of what they know.

Here, however, we start to diverge more profoundly from Bjerring and Skipper (2018)'s view. The dynamic part of their model comes not from what questions agents consider, and how their belief and knowledge states are thus actualized, but rather how many steps of reasoning they have undertaken following a specified set of inference rules. They also try to model a diamond-like dynamic operator, so that their view ends up giving us an account not of what agents *must* know after *n* steps of reasoning, but rather of what the agents *may* come to know after *n* steps of reasoning. Berto and Jago (2019, pp. 118-123) go on to criticize this view, and I think their considerations are right. Before considering them directly, and in order to better understand what's at stake, let us consider a closely related issue: that of the goal of the present views.

There is a lingering question underlying this strategy and what has been said about Jago's position: should we attempt to simply *describe* what agents know, given what they know? Or should we instead try to model what agents *ought* (rationally) to know, given what they know? That is, is our project at this point normative or descriptive? Bjerring and Skipper's (2018) account is normative, as they attempt to make explicit what minimally rational agents ought to know given what they know, and in this sense, the requirement that an agent passes the test has some more plausibility, given that what they're imagining is already the ideal situation in which an agent displays minimal rationality at all times. On the other hand, Jago's (2014) and Berto and Jago's (2019) approach is descriptive: they attempt to provide a model of what *is* epistemically possible or not for given limited agents, not what ought to be so or ought not to be so. Given the contrasting aims of both projects, it would seem that there is no conflict between them at all. What then, is it we were doing when we opposed them and came up with what is, in certain respects, an intermediate view? Answer: we were looking for a way to develop the descriptive project in a way that does not lead us to say that what is epistemically possible for given agents in certain contexts is vague, while at the same time fixing what seem to be some shortcomings of the normative version.

The puzzle for Bjerring and Skipper's solution is, then, the following: by trying to characterize a diamond-like dynamic epistemic operator, they only tell us what agents may come to know, not what agents must come to know after *n* steps of reasoning - in fact, given that what is deductible in *n* steps of reason is so varied, it would seem that there is nothing that agents must come to know after *n* steps of reasoning. But given that their view is supposedly normative, it would seem that it is precisely a box-like operator that we should be trying to model, that is, we should be trying to convey what agents must come to know after *n* steps of reasoning, given what they know. This is Berto and Jago's main criticism of Bjerring and Skipper's view.

I believe that Berto and Jago are right in their criticism of this approach. Notice, however, that if we consider the dynamic aspect of the model to not be tied with what reasoning steps the agent takes, or not only with such steps, but rather with what questions they consider then it would seem we don't face this objection. If, as we have been hypothesizing, what questions an agent considers guides their inquiry, then if we take these questions to be tied with the actions that lead to an update in an agent's knowledge state, we will no longer be in a position where an agent might come to whatever conclusions out of a great diversity of conclusions they can derive from their initial knowledge state, but only to the conclusions, if any, that are answers to the questions that the agent has considered. By limiting the space of outcomes in this way, then it becomes much more plausible to impose a restriction on what agents *must* come to know, instead of simply characterizing what they may come to know. In fact, it seems that we would arrive at the way of taking Bjerring and Skipper's test defended beforehand: perhaps what minimal rationality *requires* of an agent is that they reliably/always in ideal conditions pass the test. We could then provide an alternative normative view. As for the descriptive project, given that we treat rationality as passing the aforementioned test in certain circumstances, but not in all circumstances, we no longer need to appeal to vagueness, and it will be a determinate matter at what point agents fail to draw a trivial consequence of what they know. Whether we are able to know what that point is, however, is a different matter, and in regards to it Jago (2014) is probably right that we can't really put our finger on where exactly agents fail to draw such a trivial inference. Notice, however, that this is not to say that we have adopted here an epistemicist approach to the problem of vagueness, for we do not consider the present case to be an instance of that problem. Rather, just like for instance it might be practically impossible to ascertain exactly how many human beings have existed throughout history, but presumably there is a determinate number that is the right answer for that question, it might be impossible for us, and even the agent themselves, to find out when an epistemic oversight has occurred, even if there is a determinate answer for that question.

The impossible worlds dimension

We can now present in general lines what the impossible worlds solution here sketched for the problem of logical omniscience looks like. We start by taking thin propositions, thought of as truth-conditions, to be the content of the state of knowledge and to be sets of worlds (possible and impossible). Worlds are themselves sets of sentences of a lagadonian language, with the possible worlds corresponding to the maximal and consistent sets and the impossible worlds to all the others, so that on this view there is a world for each set of sentences. We then define deep epistemic space to be the space W of all possible and impossible worlds save for the blatantly incomplete and the blatantly inconsistent worlds - that is, the worlds that represent contradictions and the empty world (the world that does not represent anything as being the case)[104]. The space of epistemically possible worlds for a given agent at a given time is defined in terms of an accessibility relation on worlds within deep epistemic space. A given agent's knowledge is then defined as usual in terms of truth in all accessible epistemic possible worlds,

[104] In the present framework a world which does not represent anything is the same as a world containing no lagadonian language sentences. Some worlds will contain lagadonian language sentences and therefore represent something as being the case, while being empty in the sense of having no domain of objects. So, for instance a world might simply represent that everything is self-identical, without representing the existence of any object, by not containing any lagadonian language sentence where an object serves as a name for itself. Here when we say empty world we mean on the other hand a world that does not contain any lagadonian language sentences, whether they contain names or other expressions used to refer to individuals or not.

while what is compatible with a given agent's knowledge is modelled as truth in some epistemically possible worlds. The accessibility relation can then be constrained in several ways, and for each agent we want to model (in multi-agent settings) there will correspond a different accessibility relation. An example of a constraint on the accessibility relation is that the actual world should always be epistemically accessible, which amounts to the condition that knowledge should be factive. We would, finally, like to add a dynamic operator [q] defined for each agent at a given context, which tells us what happens in all circumstances where the agent, starting from a given state of knowledge (that is, with the accessibility relation on worlds defined in a specific way), considers the question q.

One crucial aspect of our approach is, then, the action of an agent considering a given question. So far, even what questions are, how agents address them, and even to an extent how their role should be characterized remains a mystery. Getting clear on all these questions is, then, a prerequisite of a plausible account that faces the problem of logical omniscience. Besides, while we might have seen what some reasons might be for incorporating questions into the present framework - that they might play a role in an agent's determination of a given world as incomplete, and that they can provide a guide for reasoning (with perhaps what questions an agent considers playing a role in how easily they can spot hidden contradictions in a given world) these haven't been made precise. The second part of the present dissertation is wholly dedicated to these issues. Once we have a clearer grasp on all of these topics, we will finally be in a position to provide a tentative solution for the problem of logical omniscience.

Summary of the first part

With the aim of dealing with some problems related to logical omniscience - among others, how to disentangle necessarily equivalent propositions; and how to model in modal approaches to knowledge agents that do not know all necessary propositions, or all logical consequences of what they know -, we first started by considering intuitive examples that motivate the claim that logical omniscience *is* indeed a problem, that is, that we should not model agents as logically omniscient, even if we aim to characterize merely implicit knowledge.

Having recognized it as a real issue, in this first part of the dissertation, we have tried to develop one way of tackling the problem of logical omniscience: that of adding worlds - impossible worlds - to the standard possible worlds framework, so that necessarily equivalent propositions come out as distinct, and that some epistemic possibilities for non-omniscient agents are genuine, metaphysical and logical impossibilities.

In order to ascertain the plausibility of such a view, we started from its metaphysical foundations and went on to consider various issues with ersatzism, on which impossible worlds' views depend on, such as Jago's (2014). Agreeing with Lewis's arguments against magical ersatzism, we reached the conclusion that the best strategy for an ersatzist is to maintain that worlds are nothing more than constructions out of bits of actuality, which are themselves concrete. We then proceeded to compare some alternative views and we opted to further develop the position according to which worlds are sets of lagadonian language sentences.

On the way, we had to deal with one of the hardest objections for ersatzists: the problem of alien objects and properties. Having considered Jago's answer to it based on Kaplan's DTHAT operator, we found it to be lacking, but at the same time that an approach based on definite descriptions could succeed. We then opted for a solution where these definite descriptions mention the worlds whose description they're a part of, so that alien objects would be represented by definite descriptions of the form: *the x such that x is P in w*. Similar considerations then applied for the case of alien properties.

Given the understanding of worlds as sets, we had then to deal with two further worries: the commitment of this last strategy to loops of proper parthood; and the controversies involving the nature of sets and their mereology. We then presented arguments for rejecting anti-symmetry, showing in the way how, accepting a view of sets as mereological sums of materially-equivalent Fine fusions, such a rejection does not have to lead to a set theory that substantially diverges from standard set-theory, for a commitment to loops of proper parthood does not entail loops of set-membership.

Having gone into detail into all these questions in how they relate to the simpler issue of the nature of possible worlds, we then showed how the framework could be expanded to include impossible worlds in a natural way. Having done so, we moved on to consider some specific issues in regards to how to construct the impossible worlds view, such as whether what counts as epistemically possible should be considered vague, or if we do better by exploring dynamic epistemic logics. We opted for the latter strategy. Finally, we considered important roles that questions might play in the framework of a dynamic approach to logical omniscience. In what follows we move on to consider various issues that arise when we try to be precise about how questions should be represented in the model, and what role they are to play.

It might seem that the questions so far considered have been dispersed and not tightly connected, but this is not so. In order to progress and to address several of the issues facing solutions for the problem of logical omniscience we had to start from the ground up and address difficulties pertaining to the builiding blocks of the theory themselves. Lewis's objection to magical ersatzism was seen to be pressing and much more general than at first might have seemed, so that it again confronted us when trying to develop certain constructivist alternatives. Further, given what was defended as a solution for the problem of alien objects and properties, it would seem that we faced a dilemma: facing Lewis's objection after all; or give up on the transitivity of the relation of proper parthood. In order to avoid both horns of the dilemma, we then had to consider various puzzles about the mereological structure of sets, so as to give a full answer to both difficulties. We ended up favouring a view which retained most of classical mereology and an intuitive conception of sets, while giving up the anti-symmetry of the relation of proper parthood. Only after considering all such matters could we then more securely claim that worlds are sets of lagadonian language sentences, which has already played an important role in the characterization of the impossible worlds solution for the problem of logical omniscience, and will continue to do so now when considering the question-sensitivity dimension of what will be the proposed solution for the problem of logical omniscience.

PART II
-
QUESTIONS AND SUBJECT MATTER

Survey of theories of subject matter

We concluded the first part of the present dissertation by considering that in order to provide an account of what is epistemically possible for a given agent in an impossible worlds framework in a way that avoids an appeal to vagueness, but that nonetheless can both describe agents' knowledge states, and ground the standards that minimal rationality imposes on agents, we needed to add question-sensitivity to our model.

So far we haven't said enough, however, about what questions are, and how we are to model them. In what follows, we adopt the strategy of putting considerations about impossible worlds for the most part on hold, and will try to show - which, as it will be seen, is a common thought - that the best theories of subject matter on offer that meet certain plausible desiderata should, in the end, accept impossible worlds. In order to do so, we start by considering various theories of subject matter on offer. We then move on to consider various issues that arise for the different conceptions, such as what general subject matters and their identity conditions are, and most importantly how they relate to the notion of a question (as we will see, while perhaps some accounts of subject matter are not objectionable for their purposes, they might break an intuitive connection between the two notions). In what follows a desideratum that we will impose on any suitable notion of question is that they be such as to be sufficiently specific to guide inquiry, so that by considering different questions, agents will be prompted to consider certain propositions and therefore prompted to reason in certain ways. After evaluating several perspectives of the nature of subject matters, we move on to consider two aboutness approaches to the problem of logical omniscience, showing how their shortcomings relate to the difficulties of trying to account for subject matters in a possible worlds framework, and how adding impossible worlds would allow both views to present more plausible ways of avoiding logical omniscience. Throughout this discussion we will also consider that despite not being closed under logical consequence, an agent's knowledge might nonetheless be restricted by weaker principles of closure, and how, besides the gains already mentioned, question-sensitivity might also help the impossible worlds theorist in this regard.

Lewisian subject matters

As it has become common, we start by considering Lewis's theory of subject matters, both because of its simplicity, and because of the place it occupies in the development of philosophical perspectives on what subject matters are, having influenced several of the existing proposals.

According to Lewis (1988b), a subject matter, what a given statement is about, is a partition of logical space induced by an equivalence relation on worlds. For instance the subject matter **how many stars there are** induces a partition in logical space in which each cell corresponds to an unstructured proposition, a set of possible worlds, that corresponds to a full and determinate answer to the question: *how many stars are there?* Lewis's theory, then, provides a direct link between subject matters and questions, and if we follow standard theories of the semantics of interrogative expressions[105] and identify questions with the set of their complete direct answers, then there is an identity between subject matters and questions on this account. Equivalently, we can think of subject matters as sets of sets of worlds, in which the sets of worlds are mutually exclusive and jointly exhaust logical space. Intuitively, all worlds in each cell agree on the number of stars, though they might diverge in several other respects. Subject matter inclusion is understood in terms of the subject matter which is included in another being a refinement of the latter's partition. For instance the subject matter **the 17th century** includes the subject matter **the 1680s** as whenever two worlds diverge in what goes on in their respective decades of 1680, they therefore also diverge in regards to what goes on in their 17th centuries. This means, then, that it is "easier" to agree - that is, more worlds do agree - in respect to the subject matter **the 1680s** than in respect to the subject matter **the 17th century**, so that each cell of the former will correspond to a union of cells of the latter. This captures an intuitive notion of subject matter inclusion, but it is important to keep in mind that the corresponding visual representation might not be the most intuitive, for the larger subject matter will be the one corresponding to the partition with smaller cells.

In spite of having various attractive features, the Lewisian view has a number of counterintuitive consequences. For instance, it is natural to try to find, for a given sentence, what the sentence is about, *the* subject matter of a sentence. In Lewis's conception of subject matters, there isn't a unique subject matter to which a given sentence corresponds, so it is not possible to determine for a given sentence what its subject matter is. Lewis ends up proposing as a least subject matter for a given statement, *P*, the subject matter **whether or not P**. This subject matter is itself included in all other subject matters in which *P* features, for this corresponds to the two-celled partition dividing logical space into the class of worlds where P is the case, and those where it is not the case, and each of these cells will always correspond to a union of cells of the refinements where *P* is involved in the subject matter. A different feature that stems directly from this last commitment is that it would seem that the least subject matter of a necessary proposition (in Lewis's framework *the* necessary proposition) should just correspond to logical space, that is, it would be the degenerate one-celled partition, which would in turn have as a consequence that every subject matter contains as a part the subject matter of the necessary proposition, which is the same as the claim that the necessary proposition is entirely about everything[106]. Lewis ends up rejecting that the one-celled partition should be considered to be a

[105] See Cross and Roelofsen (2018) for an historical and critical survey of the philosophical treatment of questions.

[106] Lewis (1988a, p. 11) defines the notion of a statement being entirely about a subject matter in the following way: a statement is entirely about a given subject matter if its truth-values supervene on the subject matter, that is, the truth-value of the statement does not depend on anything that goes beyond the subject matter. For instance, a statement is entirely about the 17th century if worlds disagreeing in regards to its truth-value belong to different cells in the corresponding partition. Given this definition, then statements about the 1680s will be entirely about the 17th century, and in general if a partition is included in another, then a statement that is entirely about the latter will also be entirely about the former. Lewis therefore accepts that if a given statement is entirely about a given subject matter, then it is entirely about all subject matters that include it. Since the one-celled partition is included in all partitions, then we would have as a result that a statement expressing a necessary propositions would be entirely about everything, in the sense of being entirely about every subject matter.

subject matter, but mainly for other reasons[107]. Lewis rejects then that statements expressing the necessary or the impossible proposition have a least subject matter, and he keeps upholding that they are about every subject matter[108].

In spite of Lewis's remarks that given that the necessary proposition does not carry any information, that is, it does not allows us to restrict the class of worlds to which the actual world might belong, this is not a problem, it is not intuitive at all, for reasons already seen in the first part of the present dissertation, that there is just one necessary and one impossible proposition, and further that every proposition's associated subject matter (that is, the partition of which they are a cell) includes a subject matter of the necessary proposition. The subject matter **the number of stars** presumably does not contain **whether 2+2=4** as a part. The move itself also strikes us as wholly artificial, for it would seem that questions such as *does Goldbach's conjecture hold?*, where there is only one fully determinate answer that holds in all worlds, are fully legitimate, and so that if we want to keep upholding an identification between questions and subject matters, that we should accept, in Lewis's framework, that the one-celled partition *does* correspond to a subject matter.

Putting aside the case of necessary propositions, in earlier parts of the present dissertation it was also seen that we might want to distinguish between necessarily equivalent propositions, which fail to be distinguished in terms of Lewisian subject matters. In fact, Lewis considers the case of relevant logicians who might want to maintain that statements should be individuated in an hyperintensional way[109], and importantly the way he proposes for someone who would take that route (which he does not endorse, for reasons already seen, such as that genuine impossible worlds would commit us to true contradictions) is to expand the domain of worlds, so that necessary propositions would not come out as corresponding to the entire class of worlds, and therefore the impossible proposition would not correspond to the empty set of worlds either (in fact there won't be such a thing as *the* impossible proposition)[110]. If we accept a Lewisian view

[107] Namely, if this partition was accepted as a subject matter, then it would be included in every subject matter, for all partitions are refinements of logical space. This would then block Lewis from defining the connection between two subject matters in terms of overlap, that is, in terms of both containing one same subject matter as part. If the one-celled partition were accepted, then every subject matter would overlap any other, so that they would always be connected. The notion of connected subject matters is itself important for Lewis (1988a) for cases where a sentence is only partially about a given subject matter.

[108] Lewis accepts that a given statement and its negation have the same least subject matters, and therefore all the same subject matters. Intuitively *P* and *Not-P* both answer the question **whether P is the case**.

[109] It follows from what has just been stated that the logical principle *ex falso quodlibet* is relevant, if we understand relevancy, as Lewis does, in terms of the connection between subject matters. Since any contradiction is impossible, its subject matter is included in every proposition's subject matter, and therefore also in that of whatever proposition stands as the conclusion of the argument. Relevant logicians that believe relevance should be an independent standard of evaluation when considering whether an argument is sound, and that on that basis want to reject *ex falso quodlibet* will, then, have to reject this feature of Lewis's view.

[110] Lewis then goes on to object that perhaps we will run into the same problem on this expanded framework, where a proposition would be necessary* if and only if it holds in every world of this expanded logical space. We have maintained, however, that impossible worlds should be open worlds, so that to each set of sentences of a lagadonian language there corresponds a world. In this setting there is no proposition that holds in every world, possible or

of what subject matters are, then a natural move to avoid some of its most unintuitive consequences is precisely to accept an ontology including impossible worlds[111].

We have found, then, in a very immediate and natural way, how impossible worlds might be helpful for the first theory of subject matter we have so far analyzed. There is, however, a neo-Lewisian approach that proposes to avoid the more problematic parts of Lewis's theory, whilst at the same time avoiding a commitment to impossible worlds. We now proceed to consider this approach.

The neo-Lewisian account of Plebani and Spolaore

Plebani and Spolaore (2021) accept most of what sets the Lewisian approach to subject matters apart from the other proposals to be explored, but introduce important tweaks with the aim of avoiding some of the most controversial features just noted.

Accepting that a theory of subject matters should attribute a specific subject matter to every sentence, and relying on the identification of subject matters with questions, they go on to consider how to relate indicative sentences with questions, which they quickly note is an issue that has seen plenty of development in Linguistics and Philosophy. They propose that the information conveyed by a sentence can be focused on in different ways, they use "Fiona is happy" as a first example, with the three different focuses being as follows:

a) FIONA is happy.
b) Fiona is HAPPY.
c) Fiona IS happy.

In a) there is a term focus, and the naturally associated question is *who is happy*?; in b) we have predicate focus, with the naturally associated question being *how is Fiona?*; and in c) the whole sentence comes into focus, and the naturally associated question is *is Fiona happy?*, which corresponds to the aforementioned binary partition that for a given proposition *P* corresponds to the subject matter **whether P is the case**. They go on to maintain that these

impossible. Given Lewis's (1986, 1988b) reservations in regards to ersatzism, he also raises several questions in regards to the details of how to extend the space of worlds, for instance what their nature would be. The first part of the present dissertation was, to a large extent, an attempt to make such an expansion precise and to give it intuitive appeal, addressing several of Lewis's concerns.

[111] Here note that since no proposition would hold in all worlds, there would be no subject matter corresponding to the one-celled partition, as this cell would have to correspond to the unique fully determinate answer to a given question, but, given how vast logical space would be, no question seems to meet this criterion. This strategy, then, would also avoid the other difficulties we have identified for Lewis's proposal in a non-artificial way. To see why no proposition would hold in all worlds, consider that since worlds correspond to arbitrary sets of sentences of the lagadonian language, which fully account for what worlds represent, then there will be very incomplete worlds, for instance containing only one lagadonian language sentence. Let A and B be two lagadonian language sentences, then there will be an impossible world {A} and an impossible world {B}, which do not have any lagadonian language sentences in common. Since they do not have any lagadonian language sentences in common, then there isn't any proposition in which both {A} and {B} are included, so that no proposition can correspond to the totality of logical space.

questions are in turn associated with the respective subject matters **happy people**; **Fiona's mood**; and **whether Fiona is happy**. In the case of a binary relational atomic proposition, instead of three focus structures, Plebani and Spolaore (2021, p. 8) accept five, using "Al is married to Tom" as an example:

a) AL is married to Tom.
b) Al is married to TOM.
c) AL is married to TOM.
d) Al is MARRIED to Tom.
e) Al IS married to Tom.

These different ways of focusing different terms correspond respectively to the subject matters **Tom's better half**, **Al's better half**, **Married couples**, **Al's relation to Tom** and **whether Al is married to Tom**. In general the strategy seems to be that whenever we have a predicate defined on *n* objects, there will be a different term focus for each subset of objects, a focus on the predicate and the sentential focus. Term focus corresponds to what objects instantiate a given property or relation; predicate focus corresponds to what a given object's qualitative and relational profile is like; and sentencial focus corresponds to the truth-value of the proposition expressed by the sentence. We can then define the subject matter of a sentence with a given focus as a partition of logical space, and the idea is that a given sentence will always have one focus or another, so that it is possible to determine on each occasion what its subject matter is[112].

Plebani and Spolaore go on to further endorse an atom-based approach to subject matter, that is, an approach for which the subject matter of complex sentences is determined by the subject matters of the atomic sentences that compose them. Just like Lewis, they accept that negation is subject matter-transparent so that the subject matter of a sentence and its negation comes out as being the same. But they further impose that the subject matter of the conjunction should be defined in terms of a function, +, that takes the subject matters of the conjuncts and outputs the subject matter of the conjunction. They go on to state, further, that since subject matters are equivalence relations, that the + function can just be understood as set intersection. This has as a result that the subject matter of two contradictions can come apart, so for instance *Fa & ~Fa* will have the same subject matter as *Fa*, whereas *Gb & ~Gb* will have the same subject matter as *Gb*. Importantly, neither of these subject matters are included in every other subject matter, so that we won't have that the impossible proposition is about everything. In regards to *ex falso quodlibet*, then, it is possible to maintain that there is a failure of relevance when an arbitrary conclusion is derived from a contradiction, as the subject matter of the

[112] Citing Carlson (2015) , Plebani and Spolaore (2021, p. 8) claim that in English "the focus is by default on the predicate in simple unary sentences (predicate focus), and on the second term in binary, relational sentences (term focus)".

contradiction will just be the subject matter of its conjunct with positive valence, which need not be connected to the subject matter of the conclusion.

In spite of its very promising features, Plebani and Spolaore's approach ends up facing the same issue that we have previously identified for Lewis's. Given that subject matters are still partitions of logical space, then what Lewis has called the librarian problem still stands: how is a librarian to know how to organize books about different branches of mathematics such as about **topology**, **algebra** and **number theory** if they all, seemingly, correspond to the one-celled partition? Further, we would still have that in the case of a sentence that expresses a necessary proposition with focus on the whole sentence just corresponds to logical space, so that its subject matter is included in every other subject matter, so that we again can't define relevance in terms of subject matter overlap.

Plebani and Spolaore admit that their approach - which they call the modest approach to subject matter - should be extended to deal with these cases, and no solution is provided in the end for the librarian's problem. Their point is simply that some instances of hyperintensionality can be dealt with in this modified framework, such as the failure of relevance in *ex falso quodlibet* inferences. Some of the hyperintensional phenomena that this neo-Lewisian approach leaves unaccounted for, however, are exactly of the sort we have been considering thus far, such as the epistemic possibility of both *Goldbach's Conjecture* and its negation. It would seem, then, that the modest approach ends up reinforcing the earlier claim that impossible worlds should be added if a Lewisian theory is to be a plausible account of subject matters. Even if only with possible worlds we can account for some forms of hyperintensionality, it would seem that impossible worlds are still unavoidable if we are to avoid the one-celled subject matter, and the counterintuitive consequences it entails.

Yablovian subject matters

Yablo's book *Aboutness* (2014) is a landmark in terms of the development of subject matter approaches to a number of different issues in contemporary Philosophy, having influenced and inspired several different approaches since. In it, Yablo maintains that what a given sentence is about is a different component of meaning which does not reduce to truth-conditions: in order to grasp the full meaning of a sentence is not only to know in what circumstances (which he takes to be possible worlds, in what he latter called a conservative approach) they are true, but also the *way* in which they are true. Yablo then goes on to identify subject matters based on this notion of how a given sentence is true at worlds.

In order to get a better grasp on Yablo's notion, consider the example of a disjunction $P \vee Q$. If it is taken to be an unstructured proposition that corresponds to a set of possible worlds, then the disjunction will just correspond to the union of P and Q. This further unstructured proposition, however, can be made true by worlds where P is true, as well as worlds where Q is true, so we can distinguish between worlds in respect to *how* $P \vee Q$ is true/false, even if their truth-value is the same. Yablo's approach, therefore, introduces notions from truthmaker semantics. The subject matter of a given sentence is defined in terms of what Yablo calls the sentence's matter - the set of its truthmakers - and its antimatter - the set of its falsitymakers -, corresponding to the unordered pair of the matter and antimatter. To know what a sentence is about is to know what makes it true and what makes it false.

Yablo's approach helps us untangle some necessarily equivalent propositions. If we accept the law of excluded middle, then *P v ~P* and *Q v ~Q* will both be necessary propositions, and while they have the same falsitymakers - none, the empty set of falsitymakers - they have different truthmakers, {P} and {~P} in the first case, and {Q} and {~Q} in the second.

Yablo then goes on to consider how to characterize content inclusion. He imposes three conditions for a given proposition, A, to include another, B:

(i) each of A's verifiers contains a verifier for B;

(ii) each of B's verifiers is contained by a verifier of A;

(iii) each of B's falsifiers is a falsifier for A.

where a verifier/falsifier just is a truthmaker/falsitymaker. To get a better grasp on it, let us consider again the case of **the 17th century** and of **the 1680s**, where the latter is intuitively part of the former. Let us suppose that a given sentence, S, has the former as its subject matter, and that S' has the latter as its subject matter. Let further *P* and *Q* be the propositions expressed by S and S' respectively. The first two conditions have it that the truthmakers for P should contain as a part the truthmakers for Q, and further that there is no truthmaker of Q that is not contained by any truthmakers of P. Intuitively, if what S says about the 17th century includes what S' says about the 1680s, then what makes what S' says true is part of what makes what S says true, and what makes what S says true contains in part what makes what S' says true. Yablo then adds the condition that everything that makes *Q* false should also make *P* false, which is also an intuitive condition: if what S says about the 17th century contains what S' says about the 1680s, then if something makes what S' says false, it thereby makes what S says false.

So far, so good. However, to turn the Yablovian approach into a specific theory, we need to say more about what truthmakers are. In order to keep a common terminology, let us refer to truthmakers as states, usually partial states whereas possible worlds are thought to be maximally consistent states - this is Fine's (2021) approach, taking a state space as a starting point and then defining the space of possible worlds from it. Yablo takes the inverse approach and defines states in terms of possible worlds, namely as non-empty sets of possible worlds, so that they are just "additional propositions". As Yablo (2014, p. 2) puts it in the Introduction to his book: "When Frost writes, *The world will end in fire or in ice*, the truth-conditional meaning of his statement is an undifferentiated set of scenarios. Its "enhanced" meaning is the same set, subdivided into fiery-end worlds and icy-end worlds". Using a different terminology, let us refer to the undifferentiated set of scenarios as a *thin proposition* and to that same set subdivided as a *thick proposition*. Intuitively, thin propositions give us a statement's truth-conditions, whereas a thick proposition gives us the truth conditions plus its subject matter. Thick propositions, then, incorporate both components of meaning and thereby represent the full content of a given statement.

Here it is of note that Yablo's approach resembles Lewis's, in which a subject matter is a set of sets of possible worlds, each corresponding to a proposition which is a full direct answer to the corresponding question. Yablo however rejects some characteristics of Lewis's conception of subject matters, holding not that the logical space should be divided by a partition induced by an equivalence relation, but rather by what he calls a division, induced by a similarity relation, so that there is overlap between the different divisions and transitivity does not hold for the similarity relation. This allows Yablo to account for statements such as *Snow is white or cold*, which can be made true by two different truthmakers at once in worlds where snow is both white and cold.

Yablo's account avoids some of the problems that Lewis's account faces, for instance the issue raised earlier that every statement should be about the impossible proposition. In Yablo's framework while any contradictions, *P & ~P* and *Q & ~Q*, do not have any truthmakers, any sets of worlds that they are true in, they nonetheless have different falsitymakers. Given that Yablo operates within a space of possible worlds, whenever a world makes *P* true, it makes *~P* false, and vice versa. So while the falsitymakers of *P &~P* and *Q&~Q* both cover logical space, they will correspond to orthogonal divisions of it, corresponding to the differing ways by which they come out false. *P & ~P*'s set of falsitymakers will be {*P*, *~P*}, which do not include or are included in any of the falsitymakers for *Q & ~Q*. Yablo can then say that *P & ~P* contain the propositions *P* and *~P*. Just like Plebani and Spolanore's modest proposal, then, Yablo's account can, therefore, explain the failure of relevance of the *ex falso quodlibet*.

It would seem, however, that Yablo's account still fails in differentiating between subject matters over which no possible worlds diverge, such as **mathematical truth** and **metaphysical truth** (Fine's (2018) example), or **topology** and **algebra**. It seems that Yablo's approach is able to distinguish between necessarily equivalent thin propositions when they are formed in different ways from other propositions, for this last difference guarantees that either their truthmakers or their falsitymakers are going to come apart. The approach however does not seem to be able to distinguish between some necessary propositions. Independently of what we take numbers to be, if we accept that truths of mathematics are the prime example of necessary truths, it would seem that propositions such as *The number two is prime* and *The number two is even* should be true for the same "reasons"[113] at all worlds, or if we accept that *All bachelors are unmarried* is true even when there are no bachelors, being true only in virtue of the definition of the terms involved, then again it would seem that we have a necessary proposition that is true for the same reason across logical space. Here notice that since truthmakers for Yablo are identified with sets of possible worlds, then if these propositions turn out to be necessary, and if they're not true in virtue of different reasons in different possible worlds, then both the propositions themselves, their truthmakers and their falsitymakers won't be able to be distinguished from one another.

[113] Here I follow Yablo's (2014) usage and by "reasons" I just mean the truthmakers of a given statement, which for Yablo are just sets of worlds.

And this seems to be inevitable if we simply start, as Yablo does, from a space of possible worlds and try to define states in terms of them.

In response to Fine (2018), Yablo (2017)[114] himself recognizes that indeed even in what Fine calls a relaxed intensional framework, not all forms of hyperintensionality can be captured. As he mentions, the point is that it is interesting to know if standard cases of hyperintensional distinctions can be captured even if one stays within an intensional framework. It would seem, then, that again we find a direct motivation for wanting to add impossible worlds right from the get go to get a better understanding of the nature of subject matters. Here note that it seems exactly because of how limited the space of states that Yablo helps himself to is that we fail to make some distinctions between contents that we wanted to make, which will bear on the problem of logical omniscience - it seems that propositions such as *Goldbach's Conjecture* and its negation will be among the cases considered in the previous paragraph, and it serves as one of the prime examples of epistemic possibilities that go beyond what is metaphysically possible. We do not have a reason to change the framework in some way or another but rather for changing it in the specific way of enriching the space of states.

Fine's situation space approach

Fine's account of subject matters involves precisely one such way of expanding the space of states. As mentioned earlier, Fine starts from the notion of a state and then proceeds to define possible worlds as a subclass of special states. Besides partial situations, Fine also allows for impossible states, states that could not possibly hold, for instance the state formed from the fusion of the states that make P and $\sim P$ true.

Fine allows for impossible states, we have been trying to show how various theories of subject matter have to gain by accepting impossible worlds, so it seems we have very quickly achieved our goal, and we could stop here the exposition and discussion of Fine's approach. Alas, matters get more complicated. Fine (2021) has presented independent arguments against impossible worlds, and against working from a framework using worlds in general, so we have to consider some of his arguments against worlds and in favour of partial states and impossible states which are not worlds.

Before we do, let us consider some details of the Finean approach to subject matters. Just like Yablo, Fine (2018) understands the subject matter of a sentence, or what he calls a sentence's bilateral content, in terms of the states that make it true and the states that make it false. Unlike Yablo, however, Fine takes a given sentence's subject matter to be the fusion of the fusion of all its verifiers with the fusion of all its falsifiers. Fine also goes on to define bilateral content, or propositions, as *ordered* pairs, with the first element being the set of exact truthmakers for the proposition, and the second element being the set of exact falsitymakers for the proposition. Exact truthmaking/falsitymaking is subsequently defined as a relation that holds between a truthmaker/falsitymaker and a proposition whenever all parts of the

[114] The dates might seem to be wrong, but Yablo's (2017) reply to Fine (2018) turned out to come out in press before Fine's initial commentary.

truthmaker/falsitymaker play a role in determining the truth/falsehood of the proposition, and nothing else but the truthmaker/falsitymaker is required to make the proposition true/false[115]. As Fine puts it, it's not just that the truthmaker necessitates or is sufficient for the truth of its corresponding proposition, but rather that it is in some sense *responsible* for its truth. Another point of difference with Yablo's approach, as already noted, is that the truthmakers and falsitymakers for a given proposition are states, but these need not be possible or complete, or constructed out of states that are both possible and complete: most states for Fine will turn out to be impossible and/or partial.

Unfortunately, Fine does not tell us more about what he takes states to be. In fact, he considers that it is an advantage of his approach that just like Kripke's work and the mathematical development of possible worlds semantics was helpful without the need to specify what possible worlds are, the same applies to his much more vast space of states. While I believe that Fine is right to claim that his new starting point might bring advantages in the treatment of several issues for which no specific metaphysics of states must be provided, just like in the case of possible worlds semantics, it seems that in order for some of its applications to be complete (for instance to the characterization of post-modal notions such as grounding, essence or explanation) an account of what they are must be provided. In this way, Fine is wrong in claiming to have an advantage over those philosophers who conceive of states in terms of sets of sentences, for instance, just like we have done in the first part of the present dissertation. While a mathematical development of a framework might not require that we provide a metaphysics for it, some of its applications might, and to conceive of states as sets of sentences is not, therefore, incompatible with Fine's goals, but complementary.

Ignoring, then, the question of what states are, we can see how Fine's approach tackles some of the difficulties we have encountered so far. Given that truthmakers/falsitymakers are not in general possible worlds, then the subject matter of a necessary proposition will not correspond to the one-celled partition of logical space. Likewise, in Fine's model different contradictions have distinct impossible exact truthmakers, so that they can be distinguished in that way. Further, Fine adheres to a popular view in the mereology of topics, according to which the topic of a given statement is the same as the topic of its negation, and according to which Boolean connectives are transparent in regards to subject matter, such that the subject matter of a disjunction of two propositions will be the same as the subject matter of the conjunction of those same propositions, which just is the fusion of the topics of the disjuncts/conjuncts. Given that the other connectives can be defined in terms of negation and either conjunction or disjunction, the same kind of reasoning will apply in order to determine the subject matter of a given complex proposition from that of its constituents.

It might be less clear, however, if Fine's view maintains a close connection between questions and subject matters. In Lewis's, Plebani and Spolaore's and Yablo's view, subject matters were constructed from sets of worlds, so that it was natural to associate a subject matter,

[115] Here note that the verifiers and falsifiers included in Fine's definition of subject matter are not simply the exact truthmakers and falsitymakers, but all the states that are parts of them as well, so that partial truthmakers and falsitymakers will be included in the fusion.

for instance a set of sets of worlds, with questions, understood as the set of all its complete and direct answers. Indeed, questions and subject matters would then be identified. In the case of Fine's model, the subject matter of a sentence is just the fusion of its truthmakers and falsitymakers. But this state, it would seem, will very often turn out to be an impossible state, since in most cases at least some of the truthmakers and falsitymakers will be incompatible. This result seems highly counterintuitive, for even if we accept impossible states and that they might be useful for a theory of subject matter, it sounds hard to believe that what sentences are about is, in general, an impossible state. This is not to say that some sentences might not be about impossible states, but it seems unnatural to claim that atomic propositions such as *Fiona is happy* are about the state in which Fiona is happy and not happy in all the possible ways for her to be so[116]. As we saw, however, bilateral propositions just are, for Fine, ordered pairs of sets of exact truthmakers and falsitymakers. But then the set of propositions that directly answer a given question will not be the same as the subject matter of a sentence that expresses a proposition which is an answer to it. It will be possible, on Fine's approach, to derive a question from a subject matter and vice-versa - a given proposition will just be the ordered pair of some of the parts of the state that corresponds to its subject matter. Even so, we started exploring theories of subject matter in order to get a better grasp on the notion of a question, for it was questions that we have argued guide inquiry. We can then ignore for present purposes what seems to be the mathematically adequate and helpful, but not intuitively justified, option of taking a subject matter to be the state formed by the fusion of a proposition's verifiers and falsifiers. We would have in this perspective, then, that questions are sets of ordered pairs of sets of exact truthmakers and falsitymakers. This, however, also sounds counterintuitive, for even if taking propositions to be ordered pairs instead of unordered pairs of sets has theoretical advantages in the setting of a truthmaking approach to content, it does not seem like there's any independent justification for imposing an order on the states that make a proposition true or false: counting the states that make it true and then those that make it false should amount to the same truth-conditions as proceeding in inverse fashion.

It would seem, then, that even if Fine's approach can effectively deal with the problems we have identified for the other theories of subject matter, it seems that it is committed to some counterintuitive characterizations of notions such as that of a proposition, and subject matter. Moreover, it seems that the primary reason for why Fine's approach is able to deal with the difficulties mentioned earlier is precisely that it enlarges the space of admissible states in order to include impossible and partial situations. This is precisely what we have done by introducing impossible worlds and to let them be inconsistent and/or incomplete sets of sentences of a lagadonian language. As we noted earlier, however, Fine criticizes impossible worlds in favour of his account. We now move on to consider these objections.

---

[116] Hawke (2017, pp. 720-1) also leverages these complaints against Fine's position, while trying to maintain that his view has greater intuitive appeal. As Hawke claims, this is not a decisive reason to reject Fine's view. However, if an alternative to Fine's view that meets all the desiderata but has greater intuitive appeal is available, then it might be preferable. As Hawke (2017, p. 720) puts it, if the only goal of a semantic theory is to accommodate for linguistic data, then Fine's view is a perfectly serviceable tool, but if we accept that such a theory should not depart for no gain from pre-theoretical intuitions, then it's possible to "break the tie" between the competing views.

What's in a name?

In his *Parts of Classes*, Lewis maintains that there is no substantive distinction between pure set theory and arithmetic based on Peano's axioms, if the notion of successor is taken as a primitive, and if, as he begrudgingly accepts, we take the notion of a singleton to be a primitive as well. Given how little conception we have of the meaning of "singleton" and "successor", Lewis claims, how can we maintain that they are different concepts? We certainly use different names for them, but then Lewis goes on to ask "what's in a name?". While in the present case, assuming that what has been said before is a plausible account of what impossible worlds are, we do have a better grasp on the concept *impossible worlds*, and in general the concept of *worlds*, so that the cases are not entirely analogous, it is still the case, as I will try to argue, that the space of worlds explored thus far is largely equivalent to Fine's own space of states, save for the fact that Fine accepts states but does not specify what they are, taking them as theoretical primitives from which he builds up his theory, but worlds have been identified with sets of sentences. In the end, to call logical space a "state space" or a "world space" will be to just use different names for one and the same thing, just under different descriptions. In order to prove this result, we first have to consider Fine's objections to impossible worlds, and provide replies that will show how there is in fact no substantive difference between the two approaches.

Fine's first objection is a version of the compositionality objection that we have previously rebutted. We have, therefore, already addressed it, following Berto and Jago's (2019) strategy, in particular by making use of the lagadonian language's special features and how diverse the space of worlds we have accepted is, which allows us to move between a thin proposition, the set of worlds in which a sentence is true, and the lagadonian language translation of that sentence - the latter will just be the unique member of the intersection of the worlds contained in the former. A different point, however, pertains to how Fine argues that in an impossible worlds' setting the clauses for the logical connectives can't be done in a uniform non-disjunctive way. He argues that the clause for negation would have to be "(*) the statement ~A is true in a world iff it is not the case that A is true in the world."[117], which leads to contradiction in impossible worlds where both A and ~A are true. But this way of conceiving the clause for negation only retains some plausibility in a setting where impossible worlds are maximal. In the present setting, all formulas are arbitrarily evaluated at impossible worlds, however, it is still possible to provide clauses for the connectives if one drops an intra-world requirement and states these clauses in terms of what lagadonian language sentences a given world contains.

The second objection has the form of a dilemma, and has to do with how many impossible worlds we should admit in our ontology. On the one hand, Fine claims, if we don't impose any restrictions on what impossible worlds we accept, then while we will get a much more fine-grained notion of entailment and proposition (understood as a set of impossible and possible worlds) than with only possible worlds, we would therefore also lose some entailments that we would want to hold. For instance it would seem that we still want *P&Q* to come out

[117] Fine (2021, p. 141).

equivalent to *Q&P*. On the other hand if we restrict impossible worlds on the basis of what's possible, then the most natural way to do so is to claim that impossible worlds respect the same entailment relations between propositions as in possible worlds, but this will entail that everything will come out true in an impossible world, and therefore that impossible worlds cannot be used to characterize a more fine-grained notion of entailment and proposition. I believe this objection was also partially addressed beforehand, when we considered the case of someone who objected to impossible worlds claiming that it is impossible that a world represents that A and does not represent that A, and therefore there must be an impossible world representing this impossibility. As we said back then, by Nolan's Principle this only entails that there is an impossible world that represents a world as representing that A and as not representing that A, which can be done without a commitment to contradictions, for only the representations of that given world will be contradictory. It is not clear as well whether, like the first objection, Fine's second challenge does not rely on a conception of impossible worlds as maximal. Further, the approach as so far developed can address Fine's concerns directly by taking different elements from both sides of the dilemma, that is: by accepting a very anarchic space of worlds, but that are nonetheless restricted, though not in the way Fine proposes. Rather than respecting the entailment relations holding between propositions at possible worlds, impossible worlds can be restricted by what different sets of combinations of individuals, properties and relations can be formed. These sets need not respect the entailment relations that possible worlds do, and further they can give us a reason to say why statements such as *Q&P* and *P&Q* are equivalent - we can either associate them with the same lagadonian language sentence, or define classes of equivalent lagadonian language sentences in regards to some permutations of its terms.

Fine's third objection starts from observing that in a possible worlds setting we can explain the necessity of A & B by the "separate necessity of A and of B"[118], while also explaining why the separate possibility of A and of B does not entail the possibility of A & B. In an impossible worlds setting, however, we cannot give an explanation of the analogous case of what is contained in a given statement. In saying A & B one is thereby saying that A and that B, but in saying A one is not thereby saying A v B, even if there is an entailment from the first to the second and from the second to the third. Fine proposes that an impossible world theorist might try to provide an answer by delimiting the space of worlds by an accessibility relation, restricting their attention to the assertively accessible worlds, which will be impossible worlds that represent that A and that B whenever they represent A & B, but which do not represent A v B whenever they represent A or represent B. This strategy, Fine goes on to claim, ends up not explaining what it sets out to explain, for it merely replicates the mystery of the difference between the "&" and "v" connectives at the semantics level. Here I believe that Fine is right, for it seems that in order to characterize the accessibility relation on worlds to yield the results we want to, we have to know in the first place what would deliver them, so that this strategy seems to be circular - we want the statement A & B to contain A and B, so we restrict our attention to those worlds that satisfy the requirements, but it is not clear what the motivation is to identify

[118] Fine (2021), p. 143.

them with the assertively accessible worlds, besides that they satisfy the requirements that we want them to satisfy. However, this need not be an objection against the introduction of impossible worlds per se, but rather a further argument in favour of what we have maintained at the end of the first part of the present dissertation: that an impossible worlds model has to gain by introducing tools from current developments in the treatment of subject matters and questions. A common way of explaining why the content of A is included in that of A & B but the content of A v B is not included in the content of A is precisely in terms of subject matter. If we adopt Yablo and Fine's talk of "ways of being true", we can then say that while A's ways of being true are parts of the ways of being true for A & B, the ways of being true for A v B are not part of the ways of being true for A, as some such ways are just ways for B to be true, which are extraneous to A's ways of being true. There is in principle no reason why an impossible worlds theorist should not adopt a view of subject matters that yields these results, as long as they're compatible with their system - and we have seen that accounts of subject matter should make distinctions for which a richer space of states is needed.

Having addressed Fine's concerns in regards to adding impossible worlds to a possible worlds' framework, we can now try to show that the differences between the current approach and Fine's are not that significant. As we have mentioned, Fine glosses over the question of what states are, and so, given that states are truthmakers and falsitymakers, what they correspond to is left open as well: all that Fine requires is that states may be partial; that they can be in mereological relations to one another; and that the space be complete, so that it also contains all the fusions that can be formed out of the states and their parts, even impossible ones. If we accept as before that sets have their members as parts and that classes can be fused together to form further classes, then we have mereological relations in our space of worlds, for instance the world w = {A, B}will contain as parts the worlds w* = {A} and w' = {B}. We also have mereological relations in the lagadonian language itself, for instance through the concatenation of operators and other lagadonian language sentences to a given starting sentence. Notice that since lagadonian language sentences are formed out of properties, relations and individuals, they naturally correspond to some ways of conceiving of facts, and facts are themselves usually taken as plausible candidates to play the role of truthmakers. If we opt for this route, then we could take the truthmakers to be the members of worlds, so that they wouldn't be the states themselves, but rather their members (alternatively we could take them to be the states and worlds to just be sets of such states). There would be no problem in taking this route, however, for we have seen that for every number of lagadonian language sentences there will be a set containing them and only them as members, so that there is a one-one correspondence between the states (the sets of lagadonian language sentences)[119] and the truthmakers (the lagadonian language sentences). We think of every world that is not maximal or consistent as impossible, so just like Fine, we characterize possible worlds as a special class of states, instead of proceeding in the opposite direction. Further, we can identify the states that Fine refers to as impossible simply with the inconsistent impossible worlds. Finally, while Fine takes the impossible states to be only

[119] Here note that if A* is a lagadonian language sentence that while {A*} is a world, {{A*}} isn't.

“fictitious fusions”, taking them to just be certain sets[120] of possible states, we give an uniform treatment of all the states - all of them are sets of lagadonian language sentences. This gives an edge to the state space as constructed so far over Fine’s.

It would seem, then, that the impossible worlds framework explored in the first part of the dissertation meets Fine’s objections, that it can be made to correspond to several features of Fine’s state space, and further that it might even be a better way of characterizing that very same state space. Fine, referring to the strategy of adopting impossible worlds as a whole, claims that once we try to move on from possible worlds by adding more worlds, we step into the “darkness of the impossible” (Fine (2021, p.149)) and fail to know exactly how and where to expand the framework, so that instead we should try to move in the opposite direction and start from states. This, however, presupposes we have no way of delimiting the construction of worlds in an independent way. If our conception of worlds is right, however, this delimitation is present from the get go, given by all the combinations of actual properties, relations and individuals, as well as some set-theoretical constructions that we can take to be the quantifiers, connectives and variables of the lagadonian language[121].

A good part of what we have called impossible worlds, Fine calls partial states or situations. Is there a substantial disagreement here? I believe there isn’t. As it was said earlier, incomplete worlds can be taken to be impossible worlds, insofar as we consider them to be candidates for complete alternative representations of actuality. If instead we thought of them simply as candidate representations of proper parts of actuality, then we could have just referred to the exact same objects instead as partial states or situations. Similarly, just as we have so far referred to the members of worlds as lagadonian language sentences, perhaps we could instead have referred to them as states of affairs, or truthmakers and falsitymakers, depending on their relation to certain propositions. While it might be that some such usages are to be preferred to others, perhaps in virtue of their different associations in the literature, it does not seem like there is anything substantial underlying what would, unless certain further requirements were imposed on the notions at stake, then be just a terminological debate. If one is not happy to call incomplete but consistent situations impossible worlds, then one can read the proposed solution for the problem of logical omniscience partially as an exercise in situation semantics. I believe nothing turns on what specific label one gives to the states after an account of their nature, structure and mereology has been given.

[120] Fine (2021), pp. 146-7 identifies several possible candidate sets to stand for these fictitious fusions. He ends up favouring a view that takes these impossible fusions to just be the sets containing as members the incompatible possible states that the fusion would be formed from, as well as the states that are parts of them. This particular choice as well as the decision to treat impossible states differently both seem to be arbitrary, and they correspond to moves that we have no need to make in the view of worlds considered beforehand: inconsistent worlds are just worlds whose members can’t correspond to propositions that are jointly true, there’s no need to treat them differently from the rest of the worlds.

[121] Not all combinations will correspond to well-formed sentences and formulas of the language, however. A natural restriction is for instance that an atomic sentence composed of an n-place predicate can only contain n objects.

Hawke's issue theory

Finally we turn to Hawke's (2017) perspective of what subject matters are, which takes into consideration various perspectives on the nature of subject matter and takes elements from them. Notably, Hawke's perspective lines up very well with the way we have conceived of worlds.

In Hawke's perspective, a subject matter is just a set of distinctions, or ways for things to be in regards to a given subject matter. Each distinction in turn is identified with a tuple containing what Hawke calls concepts, which are understood as Carnapian intensions - partial functions from worlds to individuals. Individual concepts map each world to an individual, and where the object that the concept is mapped to (if any is) is the same across all worlds, then the individual concept is taken to be a rigid designator, otherwise the individual concept corresponds to a definite description. General concepts of n-arity on the other hand map each world to a set of tuples of length n - in the special case where n = 1 we have a property, so a world is mapped to a set of objects, the ones to which the property applies. Hawke then defines a distinction to be a tuple <R, O1, …, On> where R is a general concept and O1 through On are individual concepts. A subject matter will then be a set of distinctions.

The intuition behind taking a distinction to be such a tuple is that it will divide logical space between those worlds where given objects, those referred to by the individual concepts, instantiate a given property or relation, corresponding to the general concept, and those worlds where that isn't the case. This also allows Hawke to maintain a very close relation between questions and topics, if we think of questions as the set of all its complete answers, then each such complete answer will be determined by how worlds tackle the issues raised by the distinctions: "is the world that way or not?"[122]. In fact, one can think of every distinction as being a representation in the model of exactly such an issue.

Hawke claims that contrary to Lewis's suggestion and to Fine's view, the issue-based approach does not need to make use of impossible worlds. It is not clear, however, that this is so, and Hawke does not go into detail on the reasons for why they wouldn't be necessary. Presumably, the underlying thought is that given two necessary propositions, as before, such as *2 + 2 = 4* and *Fermat's Last Theorem*, even if the distinction they induce in logical space is such that all worlds agree on it, they can nonetheless be distinguished because they are tuples of different properties, relations and individuals. Further, in Hawke's approach different contradictions can also be distinguished in this way, avoiding some of the problems identified for Lewis's approach.

Just like other proposals so far explored, however, it does not seem that Hawke's manages to avoid all problematic cases. Let us consider the example of what are, in the impossible worlds framework, distinct propositions *Goldbach's Conjecture* and *Goldbach's Conjecture v ~Goldbach's Conjecture*. Given that for Hawke the Boolean connectives are subject matter transparent, then both propositions will come out as having the same subject

[122] Hawke (2017), p. 712.

matter, that is, will correspond to the same set of distinctions, namely to the set containing only one distinction, the distinction between worlds where every even number greater than two is the sum of two prime numbers, and the worlds where that is not the case - that is, the one-celled partition. Effectively, given that one of the disjuncts of the latter proposition is a necessary proposition, and assuming for the sake of argument that *Goldbach's Conjecture* is true, then the two propositions will come out the same in an intensional framework like Hawke's. We have then that both their subject matter and their truth-conditions are the same, so it would seem that agents should have the same attitudes towards both, even if their mental contents are determined in terms of thick propositions. Yet, the latter proposition is simply a plausibly uncontroversial instance of excluded middle, and presumably widely held to be true by mathematicians working in Number Theory, whereas proving the former remains one of the most famous mathematical puzzles in the field. Given that in both cases the distinctions are made in terms of the same general and individual concepts, Hawke's theory is unable to distinguish between them. This pattern seems to generalize for all cases of necessary propositions that agents fail to know but in which the respective disjunction of them with their negations is known.

A more attentive reader will have noticed, however, that the tuples that Hawke helps himself to are exactly alike, or a way of making precise, what the lagadonian language sentences of the impossible worlds approach were taken to be, being similar to Russellian tuples, although perhaps not with the same ontological commitments (for instance both in our and in Hawke's development there hasn't been an explicit acceptance of universals). Notoriously, such an appeal allows Hawke to make hyperintensional distinctions between distinctions, that is, between partitions of logical space. These, however, were not enough to make all the distinctions we want to make in the present setting. Given how differently agents are likely to respond to *Goldbach's Conjecture* and *Goldbach's Conjecture v ~Goldbach's Conjecture*, it seems that they should correspond to different questions. We later return to this example and try to provide some insight as to how to best characterize this phenomenon. We might, however, based on the fact that even if by being sensitive to such structured entities accounts accepting possible worlds are still unable to make all the distinctions we would want to make in a theory of subject matter, have even more reason to consider adding impossible worlds to our account of subject matters. For now, let us consider two aboutness approaches to the problem of logical omniscience, see what accounts, if any, of subject matter they're based on, and then show how again impossible worlds could help make such approaches more plausible.

Critical overview of question-sensitive solutions

Having considered different ways of understanding the notion of a subject matter and how it relates to the notion of a question, we now move on to consider two proposed solutions for the problem of logical omniscience that appeal to subject matters or question-sensitivity but that do not make use of impossible worlds, and how such an addition would benefit them.

Yalcin's "Belief as question-sensitive"

Yalcin (2018) proposes to tackle some of the difficulties involving various forms of logical omniscience that a popular account of belief, which we may call map-like, faces. According to this view, championed for instance by Stalnaker (1984), an agent's beliefs are constitutively rational, in the sense that what beliefs a given agent has are inferred from their behaviour, so as to be a part of a belief-desire pair that would allow us to make sense of the agent's actions. Given that an agent has acted in certain ways, and presupposing that they have certain desires, then in order for us to understand the agent as acting rationally, we have to describe them as being in a belief state such that, were the world to be the way the agent believes it to be, then the actions the agent performs would be conducive to the fulfillment of their goals. So for instance if I wish to go to Porto from Lisbon, but I drive off to the South, then in order for my action to not count as irrational, it must be the case that I believe that I should drive South in order to reach Porto, even if this belief is in fact mistaken.

As Yalcin (2018, pp. 2-3) suggestively illustrates, on this account an agent's belief state is not given simply in terms of individual propositions, like Mentalese sentences in a belief box, and belief revision is holistic. Yalcin has us consider the case of someone who is driving and avoids a moose on the road by taking a sudden turn. In spite of the agent not consciously considering such a proposition, it seems natural to attribute to them the belief that the moose was bigger than a golfball, otherwise they wouldn't have changed trajectories in the way they did. Instead of a golfball, we could have considered other similarly small sized objects, such as a can of peas, or a squirrel. These beliefs need not be consciously present to the agent at the moment they perform the action (or at any other moment), yet they seem to carry information that was, in a certain sense, contained in what the agent believes. This picture, then, allows for a very natural understanding of the notion of implicit information or belief. In order to get at this notion of belief, in this approach the content of agents' doxastic states are modeled as sets of possible worlds, that is, thin propositions corresponding to truth-conditions. But, as we have repeatedly noted so far, this has as a consequence that agents are modeled as logically omniscient, that is, as even if only implicitly believing/knowing all that logically follows from what they believe/know, including every necessary truth.

A first step to move away from logical omniscience in a framework only containing possible worlds that has also been explored thus far is to claim that given the amount of information in their belief systems as well as their computational limitations, agents might have inconsistent beliefs, while not explicitly believing any inconsistencies, like in the example we saw earlier from Lewis (1982). The basic idea is that agents do not have access to all the information they have stored at any given moment, and so must resort only to certain fragments of their belief systems when acting. But what should we consider to be the beliefs that are available for agents at any given time, and how should we model accessible and inaccessible beliefs? Here Yalcin (2018, pp. 7-8) makes a very ingenious connection between the motivations for fragmentalism and for question-sensitivity, by identifying questions with subject matters, which are in turn understood in the Lewisian way as partitions of logical space, and by identifying the accessible or foregrounded beliefs as those sets of belief-worlds (that is, the worlds that correspond to the content of an agent's belief state) that are unions of cells of the partition, and as inaccessible or backgrounded beliefs all the other sets of belief-worlds. When we consider what beliefs are made available at any given moment for an agent, then it seems plausible to claim, if we accept that what propositions an agent is sensitive to depends on what questions they consider, that the accessible beliefs correspond to propositions which align with the distinctions between worlds induced by the question. In this way we can at the same time make intuitive sense of why fragmentation is needed, as not all propositions "speak" to all questions, as well as have a way of determining what propositions should be included in the fragments.

As Yalcin puts it, we can think of beliefs as functions from questions to answers, where the complete answers to given questions are not sensitive to every minute distinction between worlds, but only to some such distinctions. When considering different possibilities for how the world might have been, it is plausible to think that limited agents do not consider possibilities as specific as possible worlds, but rather more coarse-grained possibilities, where some distinctions are ignored. In fact, a motivation for considering possibilities to be partial states instead of possible worlds seems to follow along these lines. We can consider propositions as sets of possible worlds to correspond exactly to some such states. For instance the state of the weather being rainy in Lisbon can be thought to be the set of worlds in which the weather is rainy in Lisbon, which in the present view also corresponds to the thin proposition *the weather is rainy in Lisbon*. As Fine (2018) points out and we have noted, this notion of a partial state does not allow necessarily equivalent states to be differentiated. But then we can follow Yablo (2017) in identifying hyperintensional states with sets of sets of possible worlds. In this setting the earlier proposition would be a complete answer (let us suppose) to the question *How is the weather in Lisbon like?*, which is just the subject matter **the weather in Lisbon**, and also an hyperintensional state. But as we have seen, Yablo's strategy does not seem to allow us to distinguish between all states we would wish to distinguish - for instance we will not be able to distinguish *Goldbach's Conjecture* from the instance of excluded middle containing it and its negation. Yalcin (2018, Section 7) proposes a solution for the problem of differentiating between intensionally equivalent states, but as we will see it will still face this counterexample. Regardless of this difficulty, it is of note that accepting this view we have it that questions not only can guide inquiry but also determine the different possibilities that an agent's doxastic and epistemic state is defined on - what the different options for the world to look like for a given epistemic or doxastic agent.

An important feature that sets Yalcin's (2018) approach apart from Yablo's (2014) is that while Yalcin takes the contents of belief and knowledge states to just be sets of possible worlds, introducing complexity not in the object of belief/knowledge, but rather in the relation itself (believing ceases to be a binary relation in Yalcin's view, being in turn a relation that holds between an agent and a thin proposition relative to a given fragment and partition of logical space/question/subject matter), on approaches like Yablo's while knowledge and belief can still be considered to be binary relations between agents and the content of their doxastic and epistemic states, or ternary, if fragment-relativity is accepted, the contents themselves are not identified anymore with thin propositions, but rather with thick propositions - that is, truth-conditions plus subject matter. I take it that both general strategies are similar: stating that the content of a propositional attitude just is a set of possible worlds but that there is more to the attitude than meets the eye just seems to put a different focus than alternatively stating that the attitude is as simple as we took it to be, but that there's more to its content than meets the eye.

Yalcin's approach is ingenious and its presentation makes it attractive and very intuitive. Yet, some problems remain for this way of tackling the problem of logical omniscience. As Yalcin himself quickly takes note of, and we have noted, adding in this way question-sensitivity and fragmentation to a standard possible worlds approach does not seem to do much in regards to tackling the problem of closure under necessary equivalence, which led to counterintuitive results such as that there is only one necessary proposition, which is known by all agents in every context; as well as to cases in which it seems that agents have different attitudes towards necessarily equivalent propositions. Given his focus on the problem of closure under logical consequence, Yalcin does not develop a full solution for this issue. Still, he presents a suggestion and points in the direction of what a solution might look like.

Yalcin considers that while propositions are themselves unstructured, that nonetheless they might be associated with concepts, which are themselves understood as partitions or subject matters. So for instance Yalcin would consider that *Lavoisier is drinking water* and *Lavoisier is drinking H2O* correspond to one and the same thin proposition, yet one but not the other might be believed by a given agent that for instance possesses the concept WATER but not the concept H2O. But if the two concepts correspond to partitions of logical space, shouldn't they make the same distinctions between worlds? Yalcin's answer is that while they correspond to the same partition, one but not the other might correspond to a given subpartition embedded in a different question or subject matter. As examples Yalcin asks us to consider the subject matters **beverages** and **chemicals**, where intuitively WATER is associated with the former but not with the latter, and vice-versa for H20. Here we could model the concepts WATER and H2O respectively as the partitions corresponding to **beverages** and **chemicals**, with a distinguished subpartition, the one corresponding to **water** or **H2O**. Here it is not clear how the subpartition would be distinguished, but perhaps one way to proceed is to identify concepts with the sets containing as members a given subject matter, and one of its subpartitions as its other members. Still, Yalcin's proposal is merely tentative, so that we can leave the details of how to distinguish a given subpartition to the side. Let us consider the case of the propositions *Water is a beverage* and *H2O is a beverage*. Our linguistic practices make an utterance of "Water is a beverage" much more natural than an utterance of "H2O is a beverage", which supports Yalcin's strategy of considering WATER to be associated with a set of distinctions - those among beverages - different from those H2O is associated with - those among the chemicals. Then to know the first proposition one only needs to know what distinguishes water from other beverages, whereas to know the second one would need to be able to also distinguish H2O from other chemicals. We can then distinguish between both cases.

It is not clear however that Yalcin's suggestion will provide us with a clear way of distinguishing between concepts in all cases. Our practices are varied and it might not be clear what is needed to possess a given concept, if we understand concepts in this way. For instance in order to call attention to mechanisms underlying conspiracy theories while at the same time parodying them, it has become common to use the term "Di-hydrogen monoxide" (let us use DHMO for short) to refer to water, while stating diverse facts about water in a way as to make the receiver of the information wary about the "negative" effects that consumption of DHMO might have. Given how unusual this way of referring to water is, even in the scientific community, if we assume, if only for the sake of argument, that "DHMO" is used most often in the context of parodies of conspiracy theories, it would seem that our linguistic practices would have it that in order to believe or know that *Lavoisier is drinking DHMO*, one would need to have some beliefs or knowledge about the subject matter of conspiracy theories, or parodies. At the same time, however, it is not clear that in order to possess the concept DI-HYDROGEN MONOXIDE one needs to have any such beliefs or knowledge, for "Di-hydrogen monoxide" is a perfectly legitimate designation for water that anyone familiar with the IUPAC nomenclature rules for chemicals could understand as referring to water, despite how uncommon its usage is. But if on the other hand to possess this concept is just to know how to distinguish it from other chemicals, if no knowledge or beliefs about parodies and conspiracy theories are needed, then it's hard to explain how it is that certain agents might fail to know or believe certain propositions associated with the concept DHMO, whilst believing the very same propositions, now associated with the concept H2O. It is not clear, then, what is needed, if anything, besides what is needed to possess the concept H2O, to possess the concept DHMO. We seem to be left with no clear explanation, then, of what goes on in cases such as this one.

Here a natural thought would be to give up on distinguishing between the concepts, and to appeal directly to a metalinguistic approach, where one and the same concept might be presented in different guises, for instance. If we adopt this strategy, however, it is not clear why we haven't done so earlier in the cases of WATER and H2O, for it would seem that it was precisely to distinguish between concepts in cases where agents might have different attitudes towards necessarily equivalent propositions that Yalcin suggested we identify concepts as a partition with a distinguished subpartition. In fact, while Yalcin's example has some plausibility, it is not clear that in order to possess the concept of H2O what if any distinctions between chemicals one needs to be sensitive to: perhaps one needs to know that H2O corresponds to a molecule formed by two atoms of hydrogen and one of oxygen, but does one need to know further properties of hydrogen and oxygen in order to possess the concept of H2O? If so, which ones? How much one grasps a given concept also seems to be a matter of degree, so that it might not be clear whether there's a sufficient grasp of a concept by a given agent in order to know a given proposition. This relates to issues so far considered, for instance if Lewis (1991) was right, mathematicians may have a sufficient grasp of the concept SINGLETON for their purposes, but not sufficient when questions pertaining to the metaphysics of sets become relevant. This matter of degree also complicates matters if one wants to claim that concept possession is what's at stake in cases where agents fail to know a proposition necessarily equivalent to one already in their informational state.

It might be said in reply to the considerations put forward in the preceding paragraph that while it may be hard to individuate certain concepts, that does not invalidate Yalcin's point. One question is what concepts *are*, another is whether it is always possible to provide identity conditions for them, and how one might proceed in doing so. That what distinctions an agent must be sensitive to might not be clear in all cases does not mean that it was a mistake to identify concepts as partitions with distinguished subpartitions. I believe that such a reply would be available to Yalcin, but nonetheless that it would count in favour of his theory if it was clearer in simple cases such as H2O how to go about determining what distinctions an agent needs to be sensitive to in order to possess it. None of the difficulties so far considered are decisive, and they might be met by a development of what was a quick suggestion on Yalcin's part. Even so, there seem to be two difficulties that Yalcin's suggestion faces and that in principle no specific development of it would help alleviate.

The first difficulty has to do with how intuitively distinct concepts, associated with intuitively distinct propositions, cannot be distinguished if modelled as Yalcin proposes. This objection is in fact just a concept-level variant of one we have considered against the simple possible worlds approach to content. Consider the pair of propositions *1 * 1 = 1* and *1 / 1* = 1. Regardless of what we take numbers to be, it seems that the concepts involved in both propositions are the same, save for the former proposition involving the concept MULTIPLICATION, whereas the latter involves the concept DIVISION. If we assume that both correspond to necessary truths of arithmetic, then it would seem that they are necessarily equivalent and therefore that the only way to distinguish between them in this approach is to claim that they are associated with different concepts. We therefore need to claim that DIVISION and MULTIPLICATION are distinct concepts if we want to claim that some agents might believe or know one proposition without believing or knowing the other. In the present view, we already have the unwelcome conclusion that everything is about **division** and **multiplication**, given that their intensions are the same and correspond to logical space, which always corresponds to a union of cells of any partition - the union of all of them. But it is also not clear what subject matters we intuitively consider multiplication relative to that we do not consider division relative to. Further, Yalcin's intuitive view is that to come to possess a new concept is to come to be able to make distinctions between worlds that one was not able to make beforehand. This also applies in the case of intensionally-equivalent concepts, for then concepts might correspond to different embeddings of the same partition, so that by acquiring the concept of H2O one is then able to make distinctions between chemicals that one was not able to beforehand. But intuitively all truths involving multiplication and division are necessary, so it would seem that there are no distinctions between worlds one can do with both concepts that one cannot do with just one of them. Here to embrace the conclusion that everything is about both **division** and **multiplication** and so that two such subject matters can be picked as the partitions for which the concepts DIVISION and MULTIPLICATION respectively correspond to subpartitions would be an ad-hoc move, for given that arithmetical truths are necessary, then it seems that any such choice of partitions would be entirely arbitrary. It would seem, then, that concepts cannot just be identified with embeddings of subpartitions in partitions of logical space, at least not if we remain in a setting containing only possible worlds.

The second difficulty corresponds to an objection raised earlier to Hawke's (2017) issue based approach to subject matters, which focuses on how many necessarily equivalent propositions a concepts-based approach, as it were, can untangle. We saw that given that the Boolean connectives are transparent in regards to subject matter, the subject matters of *Goldbach's Conjecture* and *Goldbach's Conjecture v ~ (Goldbach's Conjecture)* are the same. In Hawke's (2017) view, the two propositions were also indistinguishable as they featured the same individual and general concepts. Yalcin's suggestion does not seem to be sufficiently detailed, so that we don't know how the connectives would be dealt with on his favoured development of the suggestion. But, in order to not face the same objection as Hawke's proposal, it would seem that the only option is to treat the connectives as concepts. Taking the connectives to be concepts, however, it is not clear what partition a concept like DISJUNCTION would have as its intension. Intuitively, the connectives behave in the same way in all possible worlds, namely as described by their respective truth-tables. Effectively, one could say that given that at different worlds we assign different truth-values to different atomic propositions, that the complex propositions formed out of the atomic propositions and the connectives will also be different in different possible worlds, and in this way we can distinguish between worlds in respect to the connectives. Here, however, it is not clear that we are not in a similar position as in the case of numbers. It might be thought that at given worlds a number corresponds to the extension of concepts that it does not correspond to in other worlds - to take a famous example, while in a given possible world the number of planets in the Solar System might be nine, in other possible worlds it might be eight, or some other natural number - but it is not clear that this is a case where worlds have changed in their number nine respects. Rather, it seems that the number nine is the same in both worlds, but that the changes there are between the worlds are simply in terms of the composition of the Solar System. Similarly, the connectives seem to behave in the same way in all possible worlds, even if the truth-values of the complex propositions that they can form out of atomic propositions are different from world to world. If this were right, then we would here run into the objection put forward in the preceding paragraph.

Let us assume, then, that there are indeed changes in terms of how the connectives behave from possible world to possible world. It would seem that if we understand the changes in this way, then the concepts will correspond to very fine-grained partitions. In fact, the subject matter **connectives** would correspond to the partition where each world corresponds to its own cell, as per hypothesis no two worlds render all the same propositions true/false, and therefore all worlds are distinguishable from one another in regards to how the connectives behave in them. The same applies for specific connectives, for instance if we consider the subject matter **negation**. Given that different possible worlds are distinguishable in terms of what atomic propositions are true in them, then they will also differ in terms of what negations of atomic propositions will be true in them, making every world distinguishable in terms of negation. NEGATION would then have to be identified with the partition of logical space where each world corresponds to its own cell, as embedded in a different partition. The same reasoning applies to the other connectives. But as we saw earlier, subject matter inclusion, which corresponds to the relation between a partition and its subpartitions, corresponds to a *refinement*, such that a given subject matter, S, includes another, S', if and only if S refines S', that is, if each way for things to be with respect to S implies a way for things to be with respect to S', or alternatively if each cell of S' corresponds to a union of cells of S. But if **negation** already corresponds to the subject matter where each possible world belongs to a different cell, then it is not possible to refine it further, for there are no smaller divisions we can trace in the space of possible worlds. This means that the only partition it is a subpartition of is itself (we can call it an improper subpartition). We would have, then, that NEGATION, DISJUNCTION and the other concepts associated with the connectives, which were needed to distinguish *Goldbach's Conjecture* from *Goldbach's Conjecture v ~Goldbach's Conjecture* will turn out to be the same concept. This would entail that by distinguishing the former proposition from the latter we have thereby distinguished the former from all the propositions containing Goldbach's conjecture and any connectives. But this seems to give us the wrong result, for then, assuming for the sake of argument that Goldbach's Conjecture is true, *Goldbach's Conjecture & Goldbach's Conjecture* would not be distinguishable from *Goldbach's Conjecture v ~Goldbach's Conjecture*, as both, being necessary, would correspond to logical space, both would be associated with the concepts involved in the conjecture, and, since they contain connectives, also with the most refined partition of logical space, which corresponds to the concept of the connectives. But a Mathematician who believes in the law of excluded middle will probably believe the latter, while perhaps not believing the former, which they can easily see to be equivalent to *Goldbach's Conjecture*. So even if we accept that the connectives correspond to concepts, understood as per Yalcin's suggestion, we will still identify propositions that agents might have different attitudes towards[123].

Hawke et al. (2020) raise a different objection to Yalcin's proposal. They claim that an agent's belief state might be defined on the partition containing as its only two cells *P&Q* and its negation. An agent might then come to believe *P&Q*, but since *P* does not correspond to any cell or union of cells of the partition, they might fail to believe that *P*. Hawke et al. (2020) on the other hand accept, and we will consider this option when discussing their approach, that knowledge and belief distribute over conjunction for non-omniscient agents. Accepting for the sake of argument that this is the case, then Yalcin's proposal would fail to capture an important connection between the mental contents of the limited agents it aims to model. I believe that here Yalcin could reply in one of two ways. The first would be to deny one of the premises, namely that a belief state might be defined on a partition whose only two cells are a conjunction and its negation and to claim that partitions should be constructed from atomic propositions. This would

[123] A further difficulty stems from the fact that one of the motivations for accepting that limited agents' reasoning is question-sensitive was that the possibilities that such agents can consider were not as specific as possible worlds. As Yalcin (2018, pp. 11-13) goes on to claim, the more refined a partition, the higher its encoding cost will be, as a more refined partition will entail a higher number of distinctions that the agent will have to keep in mind. But then if we accept that connectives are concepts and that they can be distinguished from possible world to possible world, we would have that, in order to possess the concept DISJUNCTION, for instance, an agent would have to be able to make distinctions between all the possible worlds. This would undermine the initial motivation, as well as impose high demands for the possession of concepts that, intuitively, are not especially hard to grasp.

prevent the counterexample directly. The second way would be to bite the bullet and accept that in some cases the agent might not *explicitly* believe part of what they explicitly believe. For instance, in the case where an agent explicitly believes R, where R corresponds to a very lengthy conjunction, the agent might not explicitly believe every conjunct of R, but might believe so only implicitly. If we consider the distinction in terms of what the agent is aware of, then perhaps such examples gain intuitive plausibility. And importantly, Yalcin's model will always have it that *P* is implicitly believed by anyone who believes that *P&Q*, even if partitions as the one proposed by Hawke et al. (2020) are accepted. Here we do not need to take a stance on whether this third objection is successful, or if what we have identified as possible replies from Yalcin would be able to meet this objection in a satisfactory way. The two objections presented here beforehand are sufficient to show that Yalcin's proposed way of dealing with the problem of logical omniscience cannot make sufficient headway in avoiding closure under necessary equivalence, in spite of some of its ingenious manouveurs.

By extending the space of worlds to include impossible worlds, we can easily distinguish between the propositions that Yalcin goes to pains to try to accommodate in a possible worlds setting. Propositions which, even if we associate concepts with propositions, we seem to not be able to distinguish. By accepting impossible worlds we can at the same time, when dealing with necessarily equivalent but distinct propositions side-step some of the problems raised in terms of how to determine the conditions that must be met in order for an agent to possess a concept, as well as difficulties in determining when two linguistic expressions correspond to the same concept or not. A full understanding of related matters might require a broader theory of the nature of concepts and how they're acquired, and these are questions worth exploring on their own. Still, at this point avoiding these issues makes it simpler to develop a solution for the problem of logical omniscience with impossible worlds.

Hawke, Özgün and Berto's "The Fundamental Problem of Logical Omniscience"

Of the perspectives so far presented, I believe that Hawke et al.'s (2020) comes the closest to providing a plausible solution for the problem of logical omniscience. Unlike Yalcin, they take the content of knowledge states to be thick propositions, given by a set of possible worlds accounting for truth-conditions and by subject matter. Further, they do not identify subject matters with Lewisian subject matters, thereby avoiding several of the difficulties for Yalcin's proposal. Like Hawke (2017) and Fine (2018), they accept that the Boolean connectives are transparent in regards to subject matter. Like Yalcin (2018) and Lewis (1982), Hawke et al. also model agents as fragmented, so that any given item of knowledge is only held relative to what they refer to as a frame of mind. These options seem to align with the most plausible results we have encountered so far when trying to model limited agents[124].

[124] Here it is of note that Hawke et al.'s (2020) objective is to tackle a specific form of the problem of logical omniscience, which they refer to as the fundamental problem. They describe this version of the problem as focusing on the notion of knowledge (by contrast to other models which wish to characterize notions such as knowability, or being in a position to know), on arbitrary agents (by contrast with approaches that focus on ideal agents or, like ours, that focus on limited agents), and which asks what the correct logic for knowledge is, such that it avoids the lower

As previously mentioned, Hawke et al. (2020, p. 731) hold that a plausible principle for knowledge is that it distributes over conjunction. So far we haven't presented reasons for supposing that to be the case. These are mainly given in terms of examples, which we can then explain in terms of subject matter inclusion. One such example is that to know that Jones and Smith are late entails knowing that Jones is late, another is that to know that Jones is horrendously late entails knowing that he is late. Intuitively, to know a conjunction one already has to know both conjuncts - it's not that the knowledge of the agent has to be extended in certain ways, as logical omniscience would seem to require, but rather that the knowledge of the conjuncts is already included in that of the conjunction. Cases such as the ones just presented serve as intuition pumps, motivating the thought that we would not ascribe to an agent knowledge of a conjunction without attributing to them knowledge of the conjuncts. We can then explain why these look so intuitive in terms of entailment and subject matter inclusion. As we've been stressing, mere entailment between two propositions does not guarantee that an agent who knows one of them will automatically come to know the other. Yet, a more plausible restriction is that knowledge should be closed under entailment that does not add new subject matter. Let us accept, as Yablo (2014) and Hawke et al. (2020) do, that contents are given by truth-conditions and subject matter. Intuitively, if A follows from a known B, that is if all A-worlds are B-worlds (the equivalence between the two notions in the specific case of a space containing only possible worlds will be important later on), and if further A does not add anything to what B is about, then the content of A is included in that of B, since both its truth-conditions and subject matter are included in that of B. I believe that this principle is plausible and tracks an important dimension of the problem of logical omniscience.

Another interesting feature of Hawke et al.'s approach is that they model knowledge non-monotonically: gaining new knowledge might lead agents to lose some knowledge they already have. They present two very suggestive examples, which follow the same pattern: an agent knows that *P*; *P* entails a given proposition *Q*; the agent comes to know that *Q* from their knowledge that *P*; but the agent attributes a higher degree of credence or for some reason has the deeply held belief that ~*Q*; which leads them to suspend their belief in *P* and thereby to lose their previous knowledge that *P*. The first example Hawke et al. present has us consider a student who has been told by her reliable tutor that Theorem 1 is correct while also being taught Theorem 2 by her teacher, who is also a reliable source of information. As it happens, Theorem 1 is correct, and the teacher made a mistake and Theorem 2 is incorrect. While studying, the student infers correctly that Theorem 1 and Theorem 2 are in fact incompatible. Given that they attribute more credibility to their teacher, they cease to believe in Theorem 1, and therefore cease to know Theorem 1, from their new knowledge that the theorems are incompatible. A different type of example also considered by Hawke et al. (2020) is that an agent might know that they do not know a given proposition, but after they come to know that proposition, they cease to know that they don't know it, by the factivity of knowledge. Cases like this are common occurrences in human agents' cognitive lives, and I find Hawke et al's arguments to be convincing, so in what follows we also take on board this feature of their model.

bound that there are no constraints on what agents know, as well as the upper bound that agents know all logic truths and what logically follows from their body of knowledge.

Finally, it is of note that this model approximates Yalcin's in considering that what an agent knows is topic-sensitive, and that it models knowledge using resources from dynamic epistemic logic. Both of these features - question-sensitivity and the dynamic aspect that a solution for the problem of logical omniscience should take - were highlighted at the end of the first part of the present dissertation as important features of the solution we wish to present. Here again, then, we are in agreement with the direction that Hawke et al.'s (2020) model follows. The main difference between our approaches, as we will now see considering objections to their model, is precisely the acceptance of impossible worlds.

Objections to Hawke, Özgün and Berto

As you may have noticed, nothing has been said thus far in regards to how Hawke et al. (2020) conceive of subject matters, save for that they do not, as Yalcin does, identify them with Lewisian subject matters, and that they accept that the Boolean connectives are subject matter transparent. It has been no oversight on our part to not mention what subject matters are on this proposal. In their approach to the problem of logical omniscience, Hawke et al. do not commit themselves to any specific view of what subject matters *are*, rather, they simply impose that they can stand in mereological relations, such that for instance the subject matter of a disjunction will be equal to the subject matter of the corresponding conjunction, which will be the same as the fusion of the subject matters of their disjuncts/conjuncts. Of course the question then becomes if there are any options for what subject matters are in a possible worlds framework that can fulfill what Hawke et al. require of them.

As we saw in the last section, both Hawke's (2017) and Fine's (2018) theories of subject matter agree with the requirements that Hawke et al. impose on the mereology of subject matters. Hawke et al. cannot, however, appeal to Fine's approach, since it relies on a space of states which includes both partial and impossible states. Even if they did, however, we have taken note of some of the issues that Fine's approach raises, which were also considered by Hawke (2017, pp. 23-26). It would seem, then, that the only perspective of the nature of subject matters so far considered that can satisfy the requirements is Hawke's. As we have seen, however, this approach will not allow us to distinguish between the contents of a given necessary truth, and the disjunction whose disjuncts are that same proposition and its negation, for both their truth conditions and their subject matters will be the same. As mentioned in the last section, this objection is considered by Hawke et al. (2020, pp. 748-749). In fact, one of the options they consider of addressing the objection is to accept impossible worlds and claim that the aforementioned propositions do not have the same truth conditions.

Another interesting option that Hawke et al. consider is to give up on the claim that for a given content to be part of another is for the latter to entail the former and for the subject matter of the former to be included in that of the latter. The case just considered would correspond to a counterexample to it. As Hawke et al. (2020, p. 748) go on to claim, knowing *Goldbach's Conjecture* is not a part of knowing *Goldbach's Conjecture v ~Goldbach's Conjecture*. This was precisely the reasoning we presented earlier in order to claim that the two propositions should be distinguished as agents might have different attitudes towards them: it seems that one can know the latter while not knowing which of the disjuncts makes it true, and therefore not knowing the former. Hawke et al. however do not develop this point further or present further reasons to consider abandoning this view of content parthood. It seems that the intuition motivating the rejection is given in terms of what an agent needs to know in order to know a given proposition. We now proceed to consider cases that help illuminate what the intuition amounts to and to introduce a notion that helps explain why the case just considered seems to yield the wrong result. But while the notion helps make sense of the intuition motivating the rejection of the aforementioned notion of content inclusion, we will see that it can't be appealed to in Hawke et al.'s framework, and that adding impossible worlds would provide an easy way of accounting for it.

The argument that the connectives are transparent in regards to subject matter seems to rely on intuitions in regards to what conversational contributions are intuitively on-topic or off-topic - for instance see, Berto et al. (unpublished) for an argument in favour of it from an intuitive assessment of whether there are contexts in which certain conversational moves but not others are on-topic. Considering the case of the connective for negation, the motivation seems to be that to say for instance that "Jane is not a lawyer" is going to be on-topic whenever "Jane is a lawyer" is, as they seem to be on-topic if we consider obvious candidates for the subject matter of each sentence, such as: **Jane's profession**; **Jane**; and **whether Jane is a lawyer**. I believe that a different intuitive line we might pursue for considering the sentences to be equivalent in regards to subject matter is that they carry information about the same questions: if we ask what Jane's profession is, even if to be told she's not a lawyer does not give us a complete answer to our question, it still allows us to rule out some alternatives to actuality - namely those where Jane is a lawyer. It seems, however, that instances of excluded middle (we ignore cases, if such there are, where excluded middle fails) do not give us any information on the question we started from, as it does not allow us to exclude any alternatives. Answering "Jane is either a lawyer or she isn't" to the question *What is Jane's profession?* seems inadequate, save perhaps for the conversational implicature that to mention the possibility of Jane being a lawyer conveys to the interlocutor that it is a significant possibility. Intuitively, this inadequacy comes from the fact that it does not give the questioner any new information in regards to Jane's profession. For this reason, while in a sense we might say that instances of excluded middle are about the same subject matter as their disjuncts, for in a sense to say that Jane either is or isn't a lawyer is to talk about Jane's profession, perhaps in a different sense they are not about what they're disjuncts are about, for it seems there is no subject matter they say something *informative* about[125].

It might be thought that the move is still ad-hoc, and that since Hawke et al. (2020) have accepted that connectives should be topic transparent, that some instances of excluded middle should not be any different. It would seem, however, that in some other cases, what we might

[125] This is only the case if we keep ignoring cases, if such there are, where excluded middle fails. If we cease to do so, then to be told that Jane is either a lawyer or she isn't is to get some information on what the world is like: it is such that *Jane is a lawyer* and/or it's negation is not indeterminate. This information is itself much more limited in respect to such subject matters as **Jane's profession**, than that conveyed by a statement of the proposition *Jane is not a lawyer*, and it might not provide new information for agents who for instance know (let us suppose) that it is a determinate matter what profession adult human beings have. We can, then, still explain why it seems inadequate to give the reply just considered.

call general and informative aboutness[126] diverge. Let us suppose that a computer has returned a number and that Jones must try to guess what number it is, and that the only information first made available to him is that the number in question is a natural number. Let us suppose further that the computer has returned the number 2734921, and that Jones might ask questions in order to guess at the number, as long as they can be answered with a "Yes" or "No", so that for instance Jones might ask whether the number the computer has returned is larger than a given number, *n*. In this way, Jones might even be able to reach the right answer in a relatively small number of steps. Let us suppose that Jones starts by taking a direct guess at what the number is, and that upon taking his guess, let us say he guessed the number is 17, Jones is given the information that the value returned by the computer is not the number 17. Intuitively, every natural number corresponds to a fully determinate answer to the question *What number has the computer returned?*, and so the proposition *The number 17 is the number returned by the computer* has the same subject matter as *The number 17 is not the number returned by the computer*, which we might take to be **the number returned by the computer**. However, given that there are infinitely many possible complete answers to the aforementioned question, knowing that the number returned by the computer is not the number 17 does not allow Jones to narrow down the number of alternatives, as there will still be infinitely many values that the computer could have returned that are compatible with his informational state.

It would seem we could think of cases like these as situations where the proposition with which the agent is confronted does not contribute to the determination of a complete answer to the question the agent is considering, that is, that it does not carry any information about the question, but the case is more complex than that. To see why, consider the alternative case where everything is as before, save for that: i) Jones is given the initial information that the number is an integer, not that it is a natural number; and ii) that the question asked by Jones is *Is the number larger than zero?* rather than taking the guess that the number the computer has returned is the number 17. Intuitively, in this case when he gets a positive answer (since the number is again, we're supposing, 2734921), Jones gains information that helps him narrow down the possible values, as he now knows it is not a negative number. Yet, it is a well-known result from set theory that there is a bijection from the set of positive integers to the set of integers, so that both sets have the same cardinality. So it would seem that in terms of what alternative answers are available, that there are as many as before Jones got an answer to his question. We would have to say, then, that *The number is larger than zero* is not informative about the subject matter **the number returned by the computer**, which seems to be the wrong result. Here we can make sense of this gain of information if we consider some facts about the relevant sets. The set of

[126] Let us say, to make the notion more precise, that a given proposition is informatively about a given subject matter when it is an answer to the corresponding question that, if true, reduces the range of possible complete answers to the question. For instance the question *Is Jane a lawyer?*, corresponding to the subject matter **whether Jane is a lawyer**, has two possible complete answers: *Jane is a lawyer* or *Jane is not a lawyer*. Given *Jane is a lawyer* the number of possible complete answers reduces to one, and likewise given *Jane is not a lawyer* the number of complete answers reduces to one. Given *Jane is either a lawyer or not*, however, we are not able to reduce the number of complete answers to the question, as both of them are compatible with the proposition. As it will soon become clear, it is not always the case that when an agent learns something about a given subject matter, that there is a reduction of the range of complete answers to a question.

integers is Dedekind-infinite, that is, it is an infinite set that has a bijection to one of its proper subsets, namely to the positive integers. So when Jones learns that the number is a positive integer, he learns that it is one of the elements of a proper subset of the integers, even if this proper subset's elements can be put into a one-one relation to the elements of the integers. We could then say that he has gained information in the sense that the set of alternative answers upon learning that the number is larger than zero is a proper subset of the set of alternative answers Jones started with.

If we adopt this strategy, then we will also have that the proposition *The number returned by the computer is not the number 17* is informative about the subject matter **the number returned by the computer**, for even if there is a bijection from the set of alternative answers after Jones comes to know the proposition, the resulting set will still be a proper subset of the set of alternative answers Jones started with. Can we distinguish somehow between both cases? Yes, if we take a proposition being informatively about a given subject matter to be a gradative notion, which makes intuitive sense as there seem to be various possible degrees between the cases so far considered of a full answer to a given question, and of an answer that provides no information in regards to a given question. We can take a proposition's degree of informativeness relative to a subject matter to be the ratio between the number of possible complete answers to the corresponding question incompatible with the proposition and the number of possible complete answers to the question[127], which will vary between 0, when an answer provides no information in regards to a given question, and 1, when it is a complete answer to the question. In the case of the proposition *The number returned by the computer is not the number 17* only one possible complete answer is excluded, which when divided by the number of complete answers to the question results in a positive infinitesimal number. We can then say that *The number is larger than zero* is more informative than *The number returned by the computer is not the number 17* as this latter proposition has an infinitesimally small degree of informativeness, whereas given that the set of positive integers and the set of integers have the same cardinality, then their corresponding infinite numbers cannot be of different orders of magnitude, and therefore their quotient cannot be infinitesimal.

A lot more would have to be said about the cases surrounding infinity, and the approach here taken revised and refined. These cases show, however, in a particularly clear way how in

[127] In the case where the number of possible complete answers is infinite, we will have a finite or infinite quantity divided by an infinite quantity, as in the cases just considered. Here we can tackle this issue in two different ways. The first one is to make use of non-standard analysis as developed by Robinson (see Button and Walsh (2018, chapter 4) for an exposition of non-standard models of arithmetic), which adds infinite and infinitesimal numbers to the reals. On this approach we have that 1 divided by an infinite number is equal to an infinitesimal, and in the case of the integers, we would have an infinite quantity divided by an infinite quantity, which is an indeterminacy in the non-standard approach (Goldblatt (1998)), save for when they are of different orders of magnitude, that is when the infinite quantities are the inverse of infinitesimals of different orders of smallness, as defined in Button and Walsh (2018, p. 84). This is the method that will be adopted here. The second way is to take the limit as a quantity approaches infinity everytime the number of complete answers in the denominator or numerator is infinite. On this approach the degree of informativeness in the first case is a positive real number arbitrarily close to 0, and when both quantities approach infinity we have again an indeterminacy. In both cases of indeterminacy, however, since the value in the numerator corresponds to a reduction of the value in the denominator, if the indeterminacy is resolvable at all, then it will still correspond to a value between 0 and 1.

certain cases the information conveyed by a proposition and its negation can be very different in regards to a given subject matter. It might be, however, that just like non-omniscient agents may not be able to have in mind all the distinctions that would distinguish every possible world from all the others, they might not be able to consider questions with an infinite number of answers[128]. Even if that was the case, which is disputable, the cases just considered are simply limit cases where the intuition that the information conveyed to an agent is in a certain sense "irrelevant" to the goals of their inquiry - in the sense that they do not contribute in a significant way to them arriving at an answer to the question they are considering. If instead we imagine a case where the agent knows that the computer has returned a value between 1 and 5000000, the number of possible complete answers to the question *What number did the computer return?* is not infinite, but is still sufficiently large that to be told that any given complete answer is not the case is to make a very small, and practically negligible, contribution to the determination of what the correct answer to the question is. In fact, even in more common examples we can easily notice how a proposition and its negation can provide more or less information in regards to a given subject matter. In the case considered beforehand of *Jane is a lawyer* and *Jane is not a lawyer*, one of them but not the other is a complete answer to the question *What is Jane's profession?*. Even if the answer *Jane is not a lawyer* is relevant, and on-topic, given how many professions Jane could have, it gives much less information than *Jane is a lawyer* in regards to the subject matter **Jane's profession**.

We can now see why Hawke et al.'s (2020, p. 748) thesis that content inclusion is not given simply by subject matter inclusion and by the content included in another to be implied by the content it is included in seemed plausible. As they claimed, it seems that knowing *Goldbach's Conjecture* is not part of knowing *Goldbach's Conjecture v ~Goldbach's Conjecture*: it seems that agents can know the latter while not knowing the former. Using the notion of a degree of informativeness, we can explain what sounds strange about the opposite position. In a possible worlds' framework, it will be the case that *p v ~p* will be part of the content of *p* where *p* is a necessary proposition, for *p v ~p* is implied by *p*, and, by the subject matter transparency of the Boolean connectives, both propositions are subject matter equivalent. This relation of inclusion does indeed strike one as odd (here note that the relation of inclusion also goes the other way, as these propositions come out equivalent in the present setting). If the considerations just put forth are correct, part of the reason why is that *p v ~p* carries less information than *p* in respect to the subject matter **whether p**. Intuitively, to know a partial answer to a question an agent does not need to know a full answer to the question at hand, whereas the opposite may be true. For instance in order to know that Jane is not a lawyer an agent does not need to know what her profession is. This is why it seems that in order to know an instance of the law of excluded middle one does not need to know which of its disjuncts is true.

[128] Here note however that distinctions between possible worlds that distinguish every world from all the others have a greater encoding cost than some distinctions with infinite cells. We can see this if we consider that worlds agreeing on what number the computer returns might nonetheless diverge in several other respects.

In fact, this seems to coincide with part of the motivation for claiming that instances of excluded middle are true simply in virtue of their form[129].

We seem, then, to be in a position to explain why rejecting the aforementioned notion of content inclusion might look plausible. Hawke et al., however, do not give a precise explanation for why we might want to reject it, and they could not explain it as we just have. In the case of necessary propositions, a notion of subject matter defined on a space of possible worlds is going to identify the complete answer to the question of whether they are the case with the completely uninformative answers, as in the case of *Goldbach's Conjecture*. Since these are identified, we cannot say that one has a higher degree of informativeness relative to the subject matter than the other. Just as before, we reach the conclusion that we need a broader space of states in order to make all the distinctions we want to make between contents. Accepting impossible worlds, we can either reject that content is given simply by truth conditions and subject matter, adding a degree of informativeness relative to a subject matter as a third condition for a given proposition's content to be a part of another, or we can keep holding that content and content inclusion is simply given in terms of truth conditions and subject matter, but that what counts as part of what an agent needs to know a given proposition is not simply given in terms of the propositions it includes, but that propositions should be ordered in terms of how informative they are in respect to a given subject matter, and that knowledge should respect this order.

In an impossible worlds setting the move corresponding to the first option is superfluous: given that we have expanded the range of truth conditions, then a necessary proposition and the disjunction of it with its negation will not be equivalent. We can then continue to identify the content of sentences with thick propositions which correspond to truth conditions (now given in terms of sets of worlds, possible or impossible) and to subject matter, and to say that knowledge is closed under entailment that does not add new subject matter. Likewise, given how diverse the resulting space of states is, including very incomplete worlds, no proposition will correspond to the totality of logical space, so that we will never run into the same problems we have in a possible worlds framework, where seemingly different propositions were identified because they all held at all possible worlds.

When asked to consider the question of whether a given proposition is the case, the agent is prompted to consider scenarios where the proposition holds, as well as scenarios in which it doesn't, so that their epistemic state is thereby restricted to worlds where either the proposition

[129] We could explain this in terms of the nature of questions. For instance if we thought of questions as partitions of logical space, composed of possible worlds (which are complete), then the cells p and $\sim p$ would jointly cover logical space, so that $p \vee \sim p$ would likewise cover logical space. We would then have that for every proposition the corresponding instance of excluded middle would be a totally uninformative answer to any question. This would be a way to motivate the claim that the law of excluded middle is true simply in virtue of its form. We do not have to assume, and we haven't, that excluded middle holds for every proposition, or simply in virtue of their form: we have simply maintained that there are some true or unproblematic instances of excluded middle, and that in those cases the truth of the disjunction might be known while which disjunct is true might not be known by a given agent, which we can characterize in terms of the corresponding instance of excluded middle being wholly uninformative about the questions its disjuncts are answers to. Here we remain neutral on what the best answer is to the problem of vagueness and other philosophical problems where the place of excluded middle has been debated, and whether the best solutions for those problems imply a rejection of the law of excluded middle.

or its negation holds, as worlds not including either of them will thereby become blatantly incomplete for the agent, insofar as they do not speak to the question they are considering[130]. In the next section we detail our favoured impossible worlds and question-sensitivity approach to the problem of logical omniscience, addressing in greater detail the issues just considered as well as others closely related to them.

Summary

Having considered various perspectives on the nature of subject matters, we have now moved on to consider two proposed solutions for the problem of logical omniscience that rely on question-sensitivity, due to Yalcin (2018) and Hawke, Özgün and Berto (2020). After considering both options, and finding strong agreement with various elements of the solution developed in Hawke, et al. (2020), we noted that even if following their tentative replies to objections, such suggestions cannot be successful in a possible worlds' framework: associating concepts with necessarily equivalent propositions does not seem to be enough to distinguish all propositions agents have different propositional attitudes towards. Namely, it seems that both positions have a hard time distinguishing between a necessary proposition and the disjunction formed from it and its negation. We then showed how in both cases adding impossible worlds to the framework would result in a simple way of avoiding the difficulties we have identified for both positions. We have, then, found more reasons to add impossible worlds to a question-sensitive approach to the problem of logical omniscience, vindicating the legitimacy of a solution that integrates elements from both the impossible worlds and subject matter approaches to the problem of logical omniscience.

[130] Both here and in what follows we ignore cases of propositions which might have an indeterminate truth-value. For instance it might be that to be told that Jones is neither bald nor non-bald is to be given information to the question of whether he is bald, as we come to know that Jones is one of the frontier cases between bald and non-bald individuals. Perhaps the way worlds could represent this indeterminacy is to contain the relevant lagadonian language sentence, and then attribute it to an indeterminate truth-value, so perhaps we could maintain that a world speaks to a given question if it contains the relevant lagadonian language sentence either as a member or as a part of a member (for instance if it contains ~*A* where A* is also a lagadonian language sentence). A full development of the present issues would have to say more about such cases, but here to further tackle issues of vagueness in a satisfactory way would take us too far. For now we make the simplifying assumption that worlds speaking to the question of **whether *p*** either represent that *p* or represent ~*p*, but as it will be seen below, there are further reasons to not take this to be in general the case.

A solution for the problem of logical omniscience

Having considered various perspectives on subject matters and two solutions for the problem of logical omniscience that rely on them, we have reached the conclusion that impossible worlds can be helpful both for the project of characterizing the very notion of a subject matter, as well as to its application in the case of logical omniscience. We now move on to propose a solution for the problem of logical omniscience that relies on both question-sensitivity and impossible worlds, as characterized in the first part of the present dissertation.

The subject matter dimension

We have not so far favoured any specific view of what subject matters are, so our first step in characterizing the present proposal is to say what we take a subject matter to be in an impossible worlds setting. Here the first point to notice is that while in an intensional setting the difficulties involved for the most part finding ways to make distinctions between subject matters that were more fine-grained than the distinctions allowed by the standard possible words approach to content, in an impossible worlds setting we are faced with the opposite difficulty. As we saw earlier, questions are to serve the purpose of guiding inquiry, so as to exclude from an agents' epistemic state impossible worlds incompatible with the information an agent has in regards to a given question, as well as worlds that do not speak to the question, that is, incomplete worlds that do not represent anything in regards to the question that the agent considers. Therefore, considering a question must make the agent exclude certain worlds from epistemic space.

In order to not introduce more complexity than is needed, let us start by considering the option of taking subject matters/questions to be partitions where each cell contains those worlds that agree on one of the complete answers to the question, and see how far we can go. We therefore inscribe ourselves in the tradition of modelling questions as sets of complete answers. A partition is a division of a set into mutually exclusive sets and that jointly add up to the initial set, that is, sets such that each element of the partitioned set is a member of one and at most one of the cells of the partition. We can also think of partitions as being induced by equivalence relations on worlds, such that two worlds are in the same cell of the partition if and only if they agree in terms of what complete answer to a question they represent[131].

Should we identify questions with all partitions of logical space, or should we impose restrictions on what partitions count as questions? We should, if questions are to limit agents' epistemic space. Intuitively, only worlds that speak to the question should be counted in.

[131] Here we ignore the complications that led Yablo to consider subject matters to be defined in terms of divisions induced by similarity relations instead of partitions induced by equivalence relations on worlds. For instance, one of the recurrent examples of a subject matter we have explored and will continue to explore, **Jane's profession**, is presumably a subject matter in which there is overlap between some of its cells, worlds in which Jane has more than one job. For this and other cases seen below, we think Yablo is right in taking subject matters to be defined in terms of a similarity relation on worlds, but henceforth we make the simplifying assumption that Jane may only have one job and that subject matters are partitions instead of divisions.

Consider for instance the question *How many stars are there?*. It would seem that worlds that do not represent anything in regards to how many stars there are should not be in any of the cells corresponding to a full answer to the question. Further, worlds that represent a merely partial answer to the question should also not be part of the partition. Suppose that a world, *w*, merely represents in regards to the number of stars that there are less than 5 million stars. Then it would be arbitrary to say that *w* belongs to any cell or another that represents the number of stars as being a number less than 5 million. Thus, instead of taking questions to be partitions of logical space, we take them to be partitions of the set of worlds representing complete answers to the questions, where each cell in the partition corresponds to one such answer. An equivalent way of defining questions is to identify the relevant partitions with the sets of sets of worlds that have certain lagadonian language sentences in common. For instance the worlds in the partition corresponding to the question *What is Jane's profession?* are worlds of the form {Pj, …} where j is Jane and P is a profession, so that if L is the property of being a lawyer and P = L, the cell corresponding to the (unstructured) proposition *Jane is a lawyer*, which is a full answer to the aforementioned question, will be the set {w, w'…} where each of the worlds in the set will be of the form {Lj, …}, that is, it will be the set of worlds containing the lagadonian language sentence Lj[132].

This restriction on what partitions count as questions or subject matters can also be given a different motivation. Lewis (1988b) discusses two ways of thinking of subject matters: the parts-based approach and the relational or partition-based approach, which he ends up favouring. As a prime example of how the former would work we have the already discussed case of the subject matter **the 17th century**, where we have said that two worlds are equivalent in regards to the relevant subject matter in case they agree on what goes on in their respective 17th centuries (for Lewis, recall, worlds are spatiotemporal entities, so that these parts correspond to temporal parts of each world, and we have then that two worlds agree in respect to a given subject matter if their relevant parts are duplicates of one another). But, as Lewis (1988b, p. 162) goes on to claim, this approach to subject matters in terms of worlds having duplicate parts is not sufficiently general, since for instance there is no spatiotemporal region where the number of stars in a given world is stored and therefore **the number of stars** is a subject matter which cannot be successfully characterized on the parts-based approach. Fine (2018) criticizes Lewis's rejection of the parts-based approach on the basis of Lewis having too-narrow a conception of what a parts-based approach can be, stemming from his conception of the nature of possible worlds. Two worlds, Fine claims, can have a part in common, but where the part is just a state, for instance the state *having 1 million stars*.

[132] As we have seen at the end of the first part of the dissertation, blatantly inconsistent worlds are ruled out from epistemic space by fiat, so that not all worlds containing Lj will be part of epistemic space to begin with. We can then either read this restriction as well as other ways in which considering a question might restrict the space of worlds as restrictions on epistemic space, from which certain worlds have already been excluded, or as restrictions on the total space of worlds, which intuitively correspond to restrictions on what worlds are admissible that only depend on what questions the agents consider, as blatantly inconsistent worlds are eliminated from epistemic space for independent reasons.

As we saw earlier, Fine (2018) works with a space of states that includes partial and impossible states and in which possible worlds are but a special case of states: maximal and consistent states. For him, then, two possible worlds are equivalent in regards to a given proposition's subject matter when they both contain a state which is a verifier or a falsifier for that proposition. Fine (2018, p. 18) then goes on to show how even this approach is not as general as the partition-based account. Consider a space of states containing only four worlds, which differ in terms of what truth-value they assign to *p*, *q*, and their negations: w, which contains a state making *p* true and a state making *q* true; w*, which contains states that make *p* and ~*q* true; w´, which contains states making ~*p* and ~*q* true; and w", which contains states making ~*p* and *q* true. Consider, then, a partition P = {{w, w'} , {w*, w"}}. Since neither w and w' nor w* and w" have any state in common, there is no way that each cell of the partition can be described in terms of the worlds in each of them containing a state in common: the parts-based approach is less general than the partition-based approach. Fine does not take this, however, to pose a challenge for the parts-based approach, but rather for the partition-based approach: it's not that the parts-based approach is not able to characterize some subject matters, but that the partition-based approach is too general, identifying more subject matters than there are.

It is not clear that Fine's example is successful in an intensional framework, for it would seem that the subject matter **whether *p if and only if q*** (accepting that for each proposition there is the subject matter of whether it is true or not) would correspond, on the partition-based approach and in a possible world's setting, exactly to the partition Fine gives as an example: it is true that *p if and only if q* whenever *p* and *q* are true or ~*p* and ~*q* are true, and false when *p* is true and ~*q* is true and vice-versa[133]. These conditions seem to correspond to the two cells of the partition just considered. Yet, Fine's objection holds when we accept more states than just possible worlds, for then *p if and only if q* will not hold in all states in which *p* and *q* are true or their negations are both true, as the states might be partial, inconsistent, or both. Then it would seem that no subject matter would correspond to such a partition, and we have thereby found another case motivating a restriction, in an impossible worlds setting, for what partitions correspond to subject matters[134].

[133] Fine can, of course, provide a subject matter for *p if and only if q*: it will correspond to the same subject matter of *p & q* and *p v q*, that is, to the fusion of the subject matters of *p* and *q*. The point here is not to vindicate a partition-based approach in contrast to a parts-based approach.

[134] In fact our earlier restriction encompasses Fine's. For Fine (2018), states work as truthmakers and falsitymakers, but nothing is said as to what their nature is. If, following a suggestion briefly noted earlier, we identify in our framework the truthmakers and falsitymakers with the lagadonian language sentences, then the requirement that worlds agree on a complete answer to the question, as we saw, amounts to them having lagadonian language sentences in common, which, given that lagadonian language sentences are identified with the truthmakers and falsitymakers, then this effectively amounts to the requirement that two worlds are equivalent in regards to some subject matter if and only if they have a state as a common part (which on our view holds given that the members of sets are parts of sets). Fine also identifies worlds with a specific kind of state, and at this point we partly depart from his view: we think of worlds as sets of states, not as themselves states, if we identify states with the truthmakers and falsitymakers (earlier we considered the different option of taking worlds to be states, and the lagadonian language sentences to be the truthmakers and falsitymakers, case in which they wouldn't be the states and our approach would then also not be strictly equivalent to Fine's). As we saw earlier, however, Fine (2021) takes impossible states to be

In the last section we considered Hawke et al.'s (2020) objection to Yalcin's (2018) position based on the principle that knowledge should distribute under conjunction, and that since there is a partition of logical space (that is, of the set of possible worlds) that has as its only cells the sets of worlds corresponding to the propositions p & q and ~(p & q), then an agent might know relative to a given subject matter that p & q but at the same time not know that p, since p does not correspond to any union of cells of the partition. We then considered that Yalcin might opt to impose as a restriction that partitions should always be defined in terms of atomic propositions, which directly blocks the objection. Here we go into more detail on how this might be done and whether there might be a further motivation for imposing a restriction on what partitions count as questions. Addressing this last issue first, it would seem that we should not simply rule out such a partition, for then it would seem that we would fail to make sense of the subject matter **whether *p & q***. So the strategy has to be more subtle than just claiming that such a partition does not correspond to a question.

We can start to see how to go about imposing the restriction if we accept a mereology of topics like that of Hawke et al. (2020), according to which the Boolean connectives are subject matter-transparent and in which for instance the subject matter of p & q and p v q are the same and correspond to the fusion of the topics of p and q. In a partition-based approach, as we saw earlier, subject matter inclusion amounts to a relation of refinement, in which a given subject matter contains another as a part if and only if it is a refinement of that subject matter, that is, if and only if the cells of the latter are unions of cells of the former. So in this approach the subject matter of p & q will correspond to the fusion of the subject matters of p and q, and will therefore be a partition with a higher resolution, that is, making more distinctions between worlds.

We have so far talked about *the* subject matter of a proposition, but as it was noted earlier, one of the features of the Lewisian conception of subject matters was precisely that a sentence could have various different subject matters. If we are interested in stating that the semantic content of a sentence is given by two components, one of which is its subject matter, then this is a disadvantage for the partition-based approach. At this point we can either follow Plebani and Spolaore (2021) or Hawke (2017) in finding a way of maintaining that there is a unique subject matter for any given statement in a partition-based approach. We take elements from both of their approaches. We said earlier that Hawke's approach comes the closest to provide a compelling account of subject matters, and it relies on a tool we have also resorted to in constructing the worlds of our perspective: the tuples of general and individual concepts that Hawke relies on correspond to the atomic lagadonian language sentences, that is, sentences only formed by lagadonian language predicates and individual terms. Hawke thinks of subject matters as sets of distinctions, where a distinction is given in terms of in what worlds the reference of certain individual concepts belongs to the extension, in that same world, of a given general concept, as well as by the worlds where this condition fails. In a possible worlds' setting a

sets of possible states, so that our view of states and of worlds as sets of states allows us to maintain a more uniform approach. But just like in the case of the view of worlds as sets of properties versus the view of worlds as properties themselves here there would also be an alternative view where worlds are just lagadonian language sentences, case in which our view would come out equivalent to Fine's in this respect. An exploration of this alternative is saved for future work.

distinction then amounts to the same as dividing the worlds into the set of worlds where a given proposition is true, and the worlds in which it is false, that is, the worlds in which its negation is true. So on Hawke's approach to know that *p & q* relative to the subject matter **whether *p & q*** is to know in what worlds both the conditions for *p* and for *q* are satisfied, and in what worlds they are not both satisfied.

Adapting Hawke's approach to our framework with impossible worlds, we can take the subject matter **whether *p & q*** to not simply correspond to a two-celled partition, but rather to a two-celled partition defined on a four-celled partition, that is, to a partition where each cell is a union of cells of the four-celled partition. Here the two cells correspond to the set of worlds that represent a way for *p & q* to be true and to the set of worlds representing a way for *p & q* to be false. Each of the four cells of the original partition will then correspond to a different way for worlds to be in regards to the subject matter, one of which is a way for *p & q* to be true, so that all of its worlds are also members of the corresponding cell of the two-celled partition, and all the others are ways for *p & q* to be false. In this regard Hawke (2017, p. 699) himself agrees with the intuitive idea behind the ways-based approach of Yablo (2014), where to know the content of a statement is not only to know in what conditions it is true or false, but also in what ways it is true or false, given by its subject matter.

We first started by claiming that when agents consider the question of whether a given proposition is the case, that their epistemic state is then defined on the set of worlds speaking to the question, that is, the set of worlds representing the proposition or its negation. But by accepting a standard view of subject matters as partitions, we then seem to have run into the same objection that Hawke et al. direct against Yalcin. We considered that to give a satisfactory reply to this objection, we should impose that the partitions associated with complex propositions should be constructed out of the partitions of the relevant atomic propositions. We then made intuitive sense of this move by considering that the partition is not simply given in terms of in what worlds certain complex propositions are true or false, but rather in terms of what worlds represent a way for complex propositions to be true and a way for them to be false. We can then generalize this move and revise the approach favoured so far: instead of taking the cells of partitions to contain worlds where certain propositions are true and worlds where they are false, they contain, rather, worlds representing ways for them to be true and ways for them to be false.

But to what correspond these ways of being true and of being false? Let us first consider the case of atomic propositions, where there is a very natural reply: the corresponding lagadonian language sentence. So for instance what makes *Jane is a lawyer* true is that Jane instantiates the property of being a lawyer, which also seems to correspond to one way for *Jane is a lawyer* to be true. Are there any other such ways? Yes. Consider for instance the lagadonian language sentence corresponding to *Jane is a responsible lawyer*. Instantiating the conjunctive property of being a responsible lawyer is a way for someone to instantiate the property of being a lawyer. Yet, truthmakers of the sort of the lagadonian sentence just alluded to contain extraneous content which does not seem to bear on the truth of the proposition, so that regardless of whether Jane is responsible or not, it is the fact that she instantiates the property of being a lawyer that makes the proposition *Jane is a lawyer* true. Following Fine (2018) and Fine and Jago (2019), let us call the lagadonian sentence that makes *Jane is a lawyer* true, that is, the lagadonian language sentence that is sufficient and responsible for the truth of the proposition but which does not contain any extraneous content, the exact truthmaker for *Jane is a lawyer*. We can then revise our approach so that all worlds in the partition contain either an exact truthmaker or an exact falsitymaker for the proposition. Given what we have said so far, it follows that the set of worlds representing an exact truthmaker for a proposition will be the set of all worlds containing a lagadonian language sentence, and that the set of worlds representing an exact falsitymaker will be the worlds containing the lagadonian language sentence formed out of concatenating a negation operator to the first lagadonian language sentence. But this just amounts to the view we have first explored: the partition will then be formed by two cells, which have as members the worlds which correspond to the truth-conditions of the associated proposition (p and $\sim p$, for a given atomic proposition p being considered). So for atomic propositions, the view that questions are partitions of the set whose members are worlds representing one of its exact truthmakers or falsitymakers is equivalent to the view that questions are partitions of the set of worlds representing that the atomic proposition is true, or that its negation is true.

Having accounted for atomic propositions, we can now move on to provide an account of the subject matter of complex propositions. Just like in Fine (2018), Hawke (2017) and Hawke et al. (2020), here we accept that the Boolean connectives should be subject matter transparent, as well as that the mereology of topics should be such that the topics of p and $\sim p$ come out as identical, as well as the topics of $p \vee q$, $p \,\&\, q$, $p \rightarrow q$ and $p \leftrightarrow q$, which correspond to the fusion of the subject matters of p and q. This fusion will be a refined partition, corresponding to the cells where p and q are true, where p and $\sim q$ are true, where $\sim p$ and q are true, and where $\sim p$ and $\sim q$ are true. The exact ways for complex propositions to be true/false are then given in terms of the exact ways for atomic propositions to be true/false. Worlds representing that p and q are true, for instance, are worlds representing an exact way for $p \,\&\, q$ to be true, while the worlds in the other cells of the partition are worlds representing exact ways for $p \,\&\, q$ to be false.

Taking this view, however, leads to some of the cases of overdetermination, which have motivated Yablo's (2014, p. 5) option for a view of subject matters as divisions instead of partitions of logical space. To see a quick example, consider a disjunction, $p \vee q$. Intuitively, the disjunction is only made false in worlds where both p and q are false, so that all the other combinations of truth-values for p and q will correspond to a way for $p \vee q$ to be true. It would seem then that both p-worlds and q-*worlds* represent exact truthmakers for $p \vee q$, so in cases where there is overlap between them, that is, worlds that represent both that p and that q, $p \vee q$ would have, in those worlds, two exact truthmakers. A natural way of accounting for such cases is to follow Yablo and keep identifying the cells of the partition with ways for propositions to be true but allow for overlap in cases of overdetermination. As it was noted earlier, while Yablo's considerations seem to be on the right track also for other reasons, here they will be ignored, simply in order to keep the presentation simpler as the main points discussed do not turn on whether subject matters are taken to be partitions or divisions of logical space. In order to meet this goal, worlds representing both that p and that q are treated as a further exact truthmaker for $p \vee q$. Intuitively, one can think of this decision in the following way: there are three ways for $p \vee q$ to be true, one where just p is true, one where just q is true, and one where both p and q are true.

Here notice that this strategy aligns with the notion of exact truthmaker just appealed to, for if *p* and *q* are individually sufficient for the truth of a proposition, then taken together they are still sufficient for the truth of that proposition, and notice that there is no extraneous content either, for even if worlds representing both that *p* and that *q* represent more than is necessary for *p v q* to be true, representing both that *p* and that *q* is not to represent anything irrelevant for the truth of *p v q*. Having addressed this concern, we arrive at a very easy way of grasping the current proposal: when complex propositions are at hand, the cells of the partition corresponding to its subject matter will correspond to the different lines in their respective truth-tables. Since the present setting includes impossible worlds, we have the guarantee that there are enough worlds representing all the required combinations of truth-values for the atomic propositions.

A worry remains for this proposal, in terms of what the worlds in the partition represent. Since the subject matter of *p & q* is just the fusion of the partitions of the sets of worlds representing exact ways for *p* to be true/false and for *q* to be true/false, then the resulting partition will contain worlds which do not represent that *p & q*. This could seem to be a counterintuitive result, as the worlds in a partition should be worlds speaking to a relevant question, and so the worlds in the partition associated with *p & q* should represent either *p & q* or *~(p & q)*, just like in the case of atomic propositions. There is indeed some pressure to further restrict the partition, but then it would seem that the subject matter of *p & q* cannot simply correspond to a fusion of the subject matters of *p* and of *q* in this approach. But as we have seen, we can take the full answers to a question to be not simply sets of worlds where certain propositions are true or their negations are true, but rather sets of worlds containing exact truthmakers/falsitymakers for certain propositions. In the case of atomic propositions, as it was seen, the two notions collapse, but in the case of complex propositions they come apart, and this result then helps explain in a different way why it is that certain worlds representing *p* and *q* but not *p & q* and vice-versa are so clearly impossible. In the first case, a world represents an exact way for a proposition to be true, without being included in that proposition's truth conditions, while in the second case a world is included in a proposition's truth-conditions but does not represent any way for that proposition to be true. Both of these options are clearly impossible as in order for a situation to count as a possible way for something to be the case, it must be a situation where what is the case is so in one way or another, and if, on the other hand, it is a possible situation where something is the case in a certain way, then it should be a situation where what holds in it is the case *simpliciter*. Worlds representing conjuncts but not a conjunction are easily seen to be incomplete because they represent a way for the conjunction to be true, while not being part of the truth-conditions of the conjunction, whereas worlds representing a conjunction but not any of its conjuncts are easily seen to be incomplete as they represent the conjunction as being the case without representing any way for it to be true. It is because of how these worlds are easily ruled out by agents when considering the question of whether the conjunction is the case that it seems plausible to impose the restriction that the worlds that are a part of the conjunction's subject-matter should be included in the conjunction's truth-conditions.

It might still be thought that this is an ad-hoc move. So here I would like to provide two further motivations for accepting that not all worlds in the subject matter of *p & q* should represent that *p & q* or that *~(p & q*). The first motivation is that following Berto et al. (unpublished), it would seem that if subject matters are to be a second independent component of content, then it should not determine nor be determined by the first component, that is, by truth-conditions. But if we imposed that *p & q* or *~(p & q)* should hold in all worlds of the relevant partition, then it would be possible to determine a proposition's truth-conditions from its subject matter, which sounds intuitively wrong: an agent can know what a statement is about without knowing under what conditions it is true. The second motivation is that we can still think in this view of the cells as full answers to the relevant questions. Even if some worlds do not represent that *p & q*, that is, if they do not contain the requisite lagadonian language sentence as a member, they still provide all the information that is needed in order to get to know that *p & q* - that is why it is so easy to see that worlds representing *p* and *q* but not *p & q* are incomplete when considering the relevant subject matter. We can then think of complete answers to a question as being sets of worlds which provide all the information that is needed to come to know what we previously took to be the complete answers, that is, cells where certain propositions and negations of other propositions hold. Thinking of it in terms of the earlier issue we have addressed of how informative an answer is to a question, then this notion of an answer still allows us to identify the complete answers to a question with the answers that give all the information needed: it seems that everything there is to know in order to know that *p & q* is true, is that both *p* and *q* are true.

Even if it seems, as we have maintained, that Hawke (2017) gets closer to providing a fully satisfactory account of subject matters, his position is not very illuminating when it comes to how we should determine with what set of distinctions a given sentence is associated. The thin proposition that the statement expresses will come with an associated distinction - that between possible worlds containing a given lagadonian language sentence, and those containing its negation, that is, those where certain objects have a given property or are related in certain ways and those where such conditions are not met - but then it is not clear what set of distinctions we should associate this distinction with. It seems that at this point Plebani and Spolaore (2021) are right in identifying various possible candidates for the subject matter of a given sentence, depending on what object or property/relation is focused on when the sentence is used to make a specific statement, which seems to change depending on conversational features. Let us imagine that Jones is organizing an event and needs to point everyone that is a lawyer and is in a lobby to a given conference room. Let us suppose that Jane is one of the people in the lobby and that Jones wants to know whether she is a lawyer, so as to know if he should direct her to the required room. In this case, it seems that Jones is not interested in getting an answer to the question of what her profession is besides insofar as it gives him information on the question that interests him: *is Jane a lawyer?*. So, for instance while if Jane replied with "I'm an accountant" she might be speaking on-topic, conveying the information that she is not a lawyer, she is giving Jones more information than is required. To see in a more suggestive way how the information she has conveyed is only relevant in the case at hand insofar as it gives information on whether she is a lawyer, suppose that everything is just as before, but that it turns out that Jane is both a lawyer and an accountant. It would seem then that the information she has given is misleading, for, putting the issue in somewhat Gricean terms, in order to make her contribution to the conversation on-topic we have to assume that when asked if she is a lawyer, by giving the information that she is an accountant Jane intends to communicate that she is not a lawyer, for otherwise she has given no information, which is readily available to her, that allows the inquirer to arrive at an answer to their question. It would seem, then, that in certain conversations the topic of *Jane is a lawyer* is not **Jane's profession** as it has been assumed to be the case, but rather **whether Jane is a lawyer**.

This example does not aim to show that a given sentence might be about multiple subject matters. It simply shows that syntactically indistinguishable statements with the same truth-conditions (in their literal interpretation) might nonetheless have different subject matters depending on features of the context in which a statement is made. Like Hawke (2017, p. 713) we leave it as an open question if a given statement's subject matter is determined pragmatically or semantically[135]. But even if we reject the Lewisian thesis that there is no subject matter that we may call *the* subject matter of a sentence, we will soon come to notice, as Yablo (2014) did, that there will be various subject matters that the sentence is about. Out of all these subject matters, Yablo picks what he calls the exact subject matter of a sentence, the set of its exact truthmakers and falsitymakers, to be *the* subject matter of the sentence. At this level of description, Yablo's view is equivalent to ours, but Yablo works within an intensional framework where the truthmakers are possible worlds and where there is no appeal to anything resembling lagadonian language sentences. In a Lewisian framework, we can also try to think of *the* subject matter of a sentence as its minimal subject, where the minimal subject matter of a sentence, S, is a subject matter, s, such that S is wholly about s and there is no subject matter s' such that s' is part of s and S is wholly about s'. A statement being wholly about a subject matter is given in terms of the statement not containing information that goes beyond the subject matter: for instance if Lewis (1988a) is right, various scientific statements are partially about **observation** even though they are not wholly about **observation**, so that the minimal subject matter of such statements might include **observation**, as no such statement will be wholly about it. On the other hand if a statement is about a given subject matter, then it also is about all the subject matters that include it. So for instance if a statement is wholly about **the 1680s**, then it is also wholly about **the 17th century**, if a statement conveys information in regards to what went on in the 1680s and nothing else, then *ipso facto* it conveys information in regards to what went on from 1601 to 1700 while not containing any information not about this larger time period. Given that we think of the subject matters of atomic propositions in terms of distinctions and construct the subject matters of the complex propositions out of the subject matters of the atomic propositions, we can define the minimal subject matter of a given statement recursively. The minimal subject matter of a statement expressing an atomic proposition, *p*, will be its two-celled partition, that is **whether *p***. By the subject matter transparency of the Boolean connectives, we have that the minimal subject matter of a statement expressing $\sim p$ will also be **whether *p***. If *p* and *q* are atoms, then the minimal subject matter of $p \vee q$ and $p \,\&\, q$ will be the set of distinctions containing the distinction between *p* and non-*p* worlds and the distinction between *q* and non-*q* worlds. Since all propositions are either atoms or can be formed out of atoms with the help of these connectives, then we can identify for every statement expressing a proposition its minimal subject matter.

[135] Similarly, we leave it open whether a statement's truth-conditions are determined semantically or pragmatically.

We have, then, identified a way to address Hawke et al.'s (2020) objection to Yalcin in a setting in which worlds are identified with sets of lagadonian language sentences. Along the way we have also provided answers for some other issues arising for theories of subject matter. Things would seem to get a bit more complicated in an impossible worlds' setting, however, as for instance we have no guarantee that a world not satisfying a given condition, that is, a world not containing a given lagadonian language sentence, will therefore include its negation. But as we saw earlier, a plausible move in an impossible worlds' framework is to take the partitions that correspond to questions to be defined on an already restricted domain of worlds - those that speak to the question at hand, that is, those that are members of cells corresponding to full answers to the question. The case of atomic propositions therefore does not add any complications in our framework, as we can still identify the minimal subject matter with the partition dividing the restricted domain of worlds into those in which the atomic proposition is true and those where its negation is true.

The restriction of an agent's epistemic state to worlds that speak to the question they are considering is not of much help in the case of complex propositions, however. For instance restricting our attention to worlds where *p & q* is the case or *~(p & q)* is the case still leaves us with impossible worlds representing that *p & q* but not representing either that *p* or that *q*, so that an agent's knowledge might not be distributed under conjunction. Here then we have to appeal to the considerations just put forth and claim that the partition corresponding to **whether *p & q*** should contain the distinctions corresponding to **whether p** and **whether *q***, such that each cell in the partition corresponds to a way for *p & q* to be true or a way for it to be false and all worlds in the former partition are worlds included in the latter two partitions, that is, all worlds representing ways for *p & q* or ~*(p & q*) to be true/false also represent *p* or its negation as well as *q* or its negation. We can make intuitive sense of this restriction as an agent who is considering whether a given conjunction is true or not seems to only be able to make progress in tackling the question if they make progress in tackling the question of whether each of its conjuncts is true. For instance it seems that in order to make progress in knowing whether *Jones and Smith are late* an agent would need to either make progress in respect to knowing whether Jones is late or make progress in respect to knowing whether Smith is late.

We can now explain why it is that an agent that knows *p & q* knows that *p* but an agent that knows that *p* might not know that *p v q*, even if the first implies the second and the second the third. In the case of *p & q* and *p*, the former implies the latter as all possible words in which *p & q* holds are also worlds in which *p* holds, and the subject matter of *p & q* includes the subject matter of *p*, as the former corresponds to a refinement of the latter. On the other hand while all possible worlds where *p* is true are also possible worlds where *p v q* is true, the subject matter of *p* does not include that of *p v q*, in fact, it's the other way around, so that the content of *p* does not include the content of *p v q*, whereas the content of *p & q* includes the content of *p*.

In the preceding paragraph it was assumed that entailment should be understood in terms of truth-preservation in all possible worlds. This is a common move in frameworks that include impossible worlds, and the idea motivating it is that impossible worlds are worlds where logic is violated (this corresponds to the second intuitive understanding of impossible worlds presented in Berto and Jago (2019, p. 31)), so that if one wants to know what entailment relations hold between what propositions, one should look only to the worlds where logic is not violated, or alternatively to worlds closed under the same logical rules as the actual world (see for instance Priest (2021))[136]. Necessarily equivalent propositions will then still entail one another, but they will not be identified, both because they have different truth-conditions, and because they might differ in terms of their subject matter.

As it was said earlier, while in a possible worlds' framework subject matters seem to play the role of allowing for more distinctions than those recognized by a view of content in terms of truth-conditions, in an impossible worlds' setting it has the opposite role of helping to impose constraints on the content of agents' epistemic states, which we have been trying to present thus far. It may be, however, that we have moved too far in the direction of constraining epistemic space. Let us consider again the case of how to distinguish between *Goldbach's Conjecture v ~Goldbach's Conjecture* and *Goldbach's Conjecture*. Assuming for the sake of argument that both are necessary truths, since entailment is given by truth preservation at all possible worlds, we have that they entail each other (though see footnote 136 for a possible way of avoiding this conclusion and block the objection). Further, we have seen that the subject matter of a disjunction just is given by the fusion of the subject matter of their disjuncts, but since the subject matter of the disjuncts is the same partition, then the subject matter of the disjunction just is the same as the subject matter of either of its disjuncts, as the fusion of an object with itself just is that same object. We have, then, that an agent who knows *Goldbach's Conjecture v ~Goldbach's Conjecture* must also know *Goldbach's Conjecture*, which was the wrong result that led us to diagnose problems with other views of subject matter. We have, then, to modify the present perspective for otherwise the addition of impossible worlds would fail precisely in accommodating the cases we introduced them to help with.

Here one might want to give up on thinking of content inclusion as consisting of a special kind of entailment, for instance entailment that does not add new subject matter. To see why, consider that in a possible worlds' setting entailment amounts to set inclusion: a given thin proposition, *p*, entails another, *q*, if and only if *p* is a subset of *q*. So in a space with only possible

---

[136] Here impossible worlds might actually help make the commonplace idea of entailment as truth-preservation in all possible worlds more plausible. The received view seems to be that unlike physical necessity, understood as truth in all of some selected accessible possible worlds (say, for instance, those having the same laws of nature as the actual world), metaphysical necessity corresponds to truth in *all* possible worlds, without restrictions. But it would seem that some such metaphysically necessary truths, such as *2 + 2 = 4* and *A surface cannot be at the same time entirely red and green* do not entail each other. The status of both propositions as metaphysical necessities depends on their content, but if logic is to be formal, then they cannot entail each other based on their content, but only based on their logical form. If one accepts this view of logic and of the nature of metaphysical necessity, then it is hard to see how both theses can be made compatible with the understanding of entailment as truth-preservation in all possible worlds, *if one does not accept impossible worlds*. By accepting impossible worlds, there will be worlds closed under classical logic for instance which are not metaphysically possible, so that there will cease to be a conflation of the notions of logical necessity and metaphysical necessity: some logically possible worlds, that is, worlds closed under a given logic (for each logical system, a different notion of logical necessity may be defined), will turn out to be metaphysically impossible, that is, will turn out to be impossible worlds, just like in the case, we propose, of epistemic possibility. Entailment can then be defined as truth-preservation in all logically possible worlds, and in this way the problematic entailment relations would be avoided.

worlds, entailment and truth-condition inclusion are identified. In an impossible worlds' framework, however, the two notions come apart: entailment is still defined in terms of a space of possible worlds, whereas truth-conditions are now much more finely distinguished. In fact, given that every set of lagadonian language sentences is a world, and complex propositions are treated as atomic in impossible worlds, truth-condition inclusion becomes a trivial notion, where the only proposition that any proposition contains as a subset is itself. Even if the notion of set-inclusion comes out arbitrary when we consider the full space of worlds, the fact that it is distinguishable from the notion of entailment is significant, for the notions of entailment that does not add new subject matter and truth-condition inclusion relative to a subject matter will also come apart. We have earlier expressed agreement with the thought that knowledge should be closed under entailment that does not add subject matter, but that was while considering two proposals restricted to a space of possible worlds, where truth-conditional inclusion is indistinguishable from entailment. Further, to accept that content inclusion amounts to inclusion of subject matter and truth-conditions is a more natural approach, for truth-conditions and subject matter have been proposed as the two components of a sentence's content.

As it was just noted, a given thin proposition including another in an impossible worlds setting is an uninteresting notion if we consider the whole space of worlds. But if instead we consider the same notion relative to a given subject matter, then it is no longer trivial. For instance, we have accepted that the subject matter of *p & q* is a partition containing as its minimal cells (that is, the cells that the cells corresponding to the full answers to the question *Is p & q the case*? are unions of) the ways for it to be true and for it to be false, so that all the worlds in the partition must also speak to the questions of **whether *p*** and **whether *q***. If we focus our attention only on the worlds in the partition, then, like in the space of possible worlds, the worlds where *p & q* holds are also worlds where both *p* and *q* hold: relative to the subject matter of *p & q*, then, all *p & q* worlds are *p* worlds and *q* worlds. We can then say why an agent who knows that *p & q* must know that *p* and that *q*: all *p & q* worlds are *p* worlds and *q* worlds, and the subject matters of *p* and of *q* are included in that of *p & q*.

Let us go back, then, to the case of *Goldbach's Conjecture* and *Goldbach's Conjecture v ~Goldbach's Conjecture*. While, let us assume, it is the case that both are true in the same possible worlds, and therefore entail one another, they are not true in the same impossible worlds, so that they can be distinguished in terms of their truth-conditions. Further, even if we relativize to any given subject matter, it won't be the case that all the *Goldbach's Conjecture v ~Goldbach's Conjecture* worlds are worlds representing *Goldbach's Conjecture* as being the case, namely because there are impossible worlds representing *~Goldbach's Conjecture* which correspond to further ways for the disjunction to be true. On our view of content inclusion, then, the content of *Goldbach's Conjecture* is not included in that of the disjunction of it and its negation, so that one might know the latter without thereby coming to know the former. This, as we saw, seems to be the right result: in some instances, we might know that either something is the case or its negation is the case, without thereby knowing which one is the case.

Relating what has just been claimed to the preceding discussion about how informative a given proposition is relative to a subject matter, we can see how our notion of content inclusion, given in terms of inclusion of subject matter and truth-conditions, respects the ordering of propositions in terms of how informative they are relative to a subject matter without this ordering being added as a third component of semantic content. Given that no proposition corresponds to the total space of worlds and that every set of lagadonian language sentences corresponds to a world, there will be no one-celled partition corresponding to a question. But as we saw earlier, it was precisely the fact that there was such a partition that led us to identify contents that we wished to distinguish, as it is only in such a partition that the truth-conditions of two distinct propositions would be identical. Since larger sets of worlds corresponding to unions of the cells in the partition will exclude less worlds in the partition, they will correspond to the restricted truth-conditions of a proposition which gives less information about the question corresponding to the partition than propositions whose restricted truth-conditions are a proper subset of its restricted truth-conditions.

So far, we have only considered how to characterize very simple subject matters in our framework. But what about subject matters that are not of the form **whether [proposition]**, such as the already considered examples of **Jane's profession** and **the 17th century**? Given what we have said earlier about how subject matters are constructed out of the subject matter of atomic propositions, then both include the minimal subject matter of every atomic proposition, so that all worlds speak to the question of whether a given atomic proposition, *p*, is true, that is, all worlds in the partition either are included in *p* or *~p*. So the subject matter **Jane's profession** will, like on possible worlds' approaches, contain only worlds representing, for each profession, either that Jane has that profession, or that Jane does not have that profession. These subject matters are, then, modelled very closely to how they were modelled in Hawke's approach, even if the framework has been greatly expanded with the inclusion of impossible worlds. We can now, then, give a plausible reply to the librarian's problem, that is, to the question of how it is that the librarian is able to categorize books in accordance with topics in which all truths are, let us suppose, necessary, such as **topology** and **algebra**. In our approach, both correspond to highly mereologically complex subject matters, including the subject matters of all propositions in both fields, which have different truth-conditions and subject matters.

One of the primary motivations for adding impossible worlds and to try to present a realistic solution for the problem of logical omniscience was how computationally limited agents were. Earlier, while considering Yalcin's (2018) approach, we also noted that this limitation should impose restrictions on how high the resolution of limited agents' and doxastic states should be, that is, on how many distinctions between possible worlds agents can be sensitive to at the same time. But if our present view is right, then subject matters as simple as **Jane's profession** and **the 17th century**, which agents seem to be able to easily consider, are also going to have a very rich mereological structure, so that it would seem real agents should have some difficulties in grasping them. We have, then, to say something about how it is that it seems agents are easily able to consider such topics.

Our reply to the present concern is based on two key thoughts: the first is that agents' grasp on a given subject matter is not an all-or-nothing matter; the second is that how hard it is to get a grasp on a given subject matter is highly dependent on what concepts an agent possesses, what they take to be the available answers to the question, their cognitive capacities and resources, as well as the complexity of the conceptual relations involved in that subject matter. In the next paragraphs we develop and motivate these two thoughts.

It is natural to attribute to an agent a better or worse grasp on a given subject matter. As a toy example, consider the case of a student who has studied for a test, but who is more comfortable with some subjects under evaluation than others. Intuitively, various examples of this sort can be constructed, but they seem often to be cases where simply an agent has more knowledge pertaining to a subject matter than pertaining to another, which we can easily account for. For instance, it might be that the agent knows the full answer to more questions included in a given subject matter (that is, that are mereological parts of the larger subject matter) than they know the answer to questions included in another subject matter; or it might be that the agent is able to give more informative answers to the questions included in one subject matter than to those included in the other. These cases, however, do not correspond to the first thought just mentioned, as the way they can be explained away does not relate to the question just considered of certain subject matters corresponding to partitions that are too fine-grained for human agents to be able to consider.

The claim that an agent's grasp on a subject matter can be more or less limited amounts to the claim that when a question itself encompasses too many distinctions, an agent might not be able to consider the question in its full generality, but only partial questions that approach to various degrees the complete question. Here the grasp in question is not given in terms of to what extent an agent is able to give answers to the questions included in a given question, but rather in terms of being able to consider the distinctions required to give the answers to it. It corresponds to the difference between the extent to which an agent is able to answer a question, and whether they are in a position to answer it in the first place. Going back to the example of a student who is more comfortable with certain topics under evaluation in a test than others, it might be that besides having varying knowledge on those topics, in the ways already described, the student might have a more or less limited grasp on what the possible answers are to the questions. Here a helpful example is to consider the subject matter **Jane's profession** again, and that the correct full answer to the question of what Jane's profession is unusual, such as *Jane is an art therapist*. An agent who does not know that art therapist is a profession - for instance they may know of different forms of therapy, as well as of various professions that involve art in some form, but they do not know that there is a form of therapy that uses art and therapists specializing in it - will then not be in a position to give a correct answer to the question. This failure does not seem to be explicable in terms of the agent lacking knowledge of what cell of the partition corresponds to the correct full answer to the question. Rather, it seems that instead the agent fails to have in mind the partition that corresponds to the question, namely because they are not sensitive to some of its distinctions - not knowing that art therapist is a profession, the agent then is not able to consider the distinction between worlds where Jane is an art therapist and worlds where she isn't, or at least they fail to consider that distinction in relation to the specific question of what her profession is.

The case just described does not seem to be a rare occurrence, and it might be that for various questions agents consider in their day-to-day life, they do not consider the full question they take themselves to be considering. For instance, in the case of very general subject matters such as **biology** in order to be able to fully grasp it, an agent's epistemic state would need to be sensitive to every relevant distinction, so that one would need to know everything in regards to how the world may be in regards to biology, even if one does not know which of these ways correspond to how things actually stand relative to the subject matter at hand. It is implausible, however, that any agent already has this knowledge. For instance, Darwin ignored a lot of what has since been discovered about genetics, so that while his theory of natural selection is widely held to have been correct in various respects, it does not speak to questions of genetics relevant in biology as developed nowadays, such as how certain genes are passed on from a progenitor to its offspring. By coming to know that genetics plays a role in evolution, we thereby come to have a better understanding of the subject matter **biology**, even if we have a very limited knowledge of what is true about genetics. It is plausible that in regards to general and evolving topics such as **biology** our ignorance is not simply given in terms of determining what way the actual world is relative to it, but also what all the ways for worlds to be in respect to **biology** even are.

We can then define a notion of a partial question in terms of a question to which certain distinctions were subtracted, so that partial questions will be parts of their corresponding full questions. This is intuitive: to be able to consider a question in its full generality is to therefore be able to consider all of its partial domains of application. With the notion of a partial question we can then say that an agent has a grasp on a question to a greater or lesser extent depending on how much needs to be subtracted from a given question in order to get at the set of distinctions the agent's epistemic state is sensitive to. This move, however, presupposes that we move from the full question to the partial question, and define the latter notion in terms of the former, and therefore presupposes that we're able to tell what the given full question is that the agent's epistemic state would ideally be sensitive to. In cases such as the one just mentioned of **Jane's profession**, it seems clear what the full question amounts to, even if it were the case that no agent actually knew what all the professions are, but it might be that there are instances where it is not clear from what an agent's epistemic state is sensitive to what the question they are considering is a partial question of. For instance, it might be that an agent is considering questions at the border of two academic fields of inquiry but where their thoughts are not clear enough so as for it to be possible to determine whether they're considering a question pertaining to one field or the other. In such cases we can say one of three things: that it might be that there is no fact of the matter as to what full question the agent is aiming to answer; that what full question the agent is considering is ambiguous; or that the question that the agent is striving to give an answer to is itself partial, ignoring certain distinctions[137]. For different cases, one or the other of these three claims will be more plausible than the others, so that for instance in the case where an agent does not have a clear conception of the subject matter at hand, it might be that there is no fact of the matter as to what specific full question they are considering, but perhaps in certain cases an agent might purposefully be ignoring certain options, for instance because they attribute to them a very low subjective probability, so that perhaps the question they are considering is itself partial and may even be so regarded by the agent.

Now that we have motivated the notion of agents having a better or worse grasp on a certain question, we move on to briefly consider ways in which agents might be constrained in their ability to consider the requisite questions. One such case was already noted: the agent might not know that a given option is an alternative answer for a question they consider. A closely related kind of case is of agents who do not possess the required concepts to be sensitive to some of the distinctions at hand. As an example, we can consider a small child who cannot yet differentiate between the various specific professions falling under the umbrella term "scientist", but who wishes to become a scientist when they grow up. A third kind of case illuminates why for instance certain branches of pure mathematics are so hard to get a grasp on. It is common for

[137] This last restriction is not psychological like the rest, but it might be the case (as is considered at the end of the paragraph) that the reason why the agent considers the partial question is itself psychological.

statements in this and other disciplines to express highly complex relations between certain concepts. It might be, then, that an agent fails to consider certain options as answers to a given question, not because they do not possess the concepts involved, or ignore the fact that it may be an option, but rather because they fail to put together the concepts at hand in the required way, as doing so might encompass a high computational cost (as well as costs in terms of other resources, such as time). It might be that there are other ways in which agents fail to be sensitive to all the distinctions included in a question they wished to consider, and therefore this should not be taken to be an exhaustive list of reasons for why an agent might fail to fully consider a question.

The dynamic aspect

Having addressed several issues in regards to how subject matters should be characterized in an impossible worlds setting, avoiding the pitfalls previously identified for other perspectives, we now turn to consider how they play a role at large in the present approach to the problem of logical omniscience.

As it has been maintained earlier, an agent's epistemic state changes depending on what questions they consider, similarly to how in Bjerring and Skipper's (2018) model it changes depending on what reasoning steps they take. Their approach, however, ended up running into the problem of not being able to characterize a box-like dynamic operator, which would tell us what holds in all the worlds of the updated model, where an agent has taken the action of taking a certain number of reasoning steps using the relevant rules for reasoning. As it was seen, it is precisely such a box-like operator that we wish to characterize, telling us, in the descriptive approach, what the agent comes to know after considering a given question, and in the normative approach what a (minimally rational) agent ought to come to know after considering it. Only with such a box-like operator can we impose restrictions in an impossible worlds setting on what an agent knows, so that it won't be the case that for minimally rational agents that the only closure principle on knowledge is that an agent knows that *p* if and only if an agent knows that *p* (see Hawke et al. (2020, p. 731) for a succinct and clear argument for why the logic for explicit knowledge should be more constrained than this).

In this regard our approach fares better than Bjerring and Skipper's. Since questions are being understood as certain special partitions of sets, then after considering a given question an agent's epistemic state will be defined on the set of worlds the question is a partition of, that is, the set of worlds that, for a given atomic proposition *p*, contains the worlds representing *p* and representing ~*p*. These are, intuitively, the worlds that speak to the question at hand. Similarly, when agents consider complex propositions, defined in terms of its ways of being true[138], they consider all the ways in which they are true, and so their epistemic state is then made relative to

[138] For atomic propositions, there is only one way for them to be true and false, which correspond to the two cells in the partition, so that there are as many ways for them to be true, as the truth-values they may take in a world. In the case of complex propositions, however, as it has been seen, there may be more ways for a proposition to be true and false than just two, as for instance there seem to be three ways for *p* & *q* to be false: *p* being false and *q* true; *p* being true and *q* false; and both being false. Subject matters are then defined more broadly in terms of ways of being true and ways of being false.

worlds that speak to the question, as well as to the questions that need to be determined in order to determine a way for the complex proposition to be true/false. The agents thereby rule out as epistemically impossible worlds that represent neither that *p* nor that *~p*, on the count of such worlds being blatantly incomplete after they consider the question of whether *p* is the case, in the case of atomic propositions. And likewise, they rule out worlds that do not speak to the question of whether *p* is the case when they consider whether certain complex propositions that include *p* are the case, such as the question of whether *p & q* is the case.

Appealing to our notion of a partial and full question, we can also model both descriptively what agents know, as well as a notion of what they ought to know. As it was just seen, agents may fail to consider questions in their full generality for a variety of different reasons, such as them not possessing certain concepts, not considering certain options to be answers to the question they're considering, or that the conceptual relations they need to be sensitive to in order to be able to distinguish a given cell in a partition having a high computational cost. We can then define both a box-like dynamic epistemic operator where the action the agent performs is to consider a given question *in full* as well as a box-like dynamic epistemic operator where the question considered is the partial question that the agents' epistemic state is defined on. In the first case we model the upper limit of what an agent is able to rule out from epistemic space by considering a given question[139], whereas in the second case we model what an agent actually rules out from epistemic space by considering a given question.

Brief recap of the proposed solution

Before moving on to consider remaining objections for the view so far presented, let us briefly go over some of its main features. First, knowledge is modeled in a space of worlds to which there were added impossible worlds. All such worlds are themselves sets of lagadonian language sentences, which are themselves constituted by actual individuals, properties, relations and set theoretical constructions. To each set of sentences there corresponds a world, with possible worlds being the limit case of maximal and consistent sets. Some worlds, however, are not epistemically possible for any agents, as they are either blatantly inconsistent or blatantly incomplete and therefore easily seen to not be a possible alternative to actuality (which must be both consistent and maximal, as all possible worlds are). The blatantly inconsistent worlds can be taken to be worlds representing a proposition and its negation, as well as worlds representing their conjunction, while the blatantly incomplete world is the empty set of sentences. With both notions in place, we can then introduce a notion of a rank, which tells us how inconsistent or incomplete a world is. In the case of inconsistency, we followed Berto and Jago (2019) and accepted that subtly inconsistent worlds can be linked to blatantly inconsistent worlds by proofs, where we move from a world where the premises but not the conclusion of an argument are true, to a world where also the conclusion is true. The rank of a world is then given in terms of the

[139] Here note that by containing more distinctions than any of its corresponding partial questions, a full question will impose more restrictions on epistemic space. Going back to the example of Jane's profession, it might be that for an agent some worlds in the partial question's partition do not represent anything in regards to Jane being an art therapist or not, but all worlds in the full question's partition represent either that Jane is an art therapist, or that she isn't.

number of steps needed to, in this way, arrive at a blatantly inconsistent world, so that the higher the rank of a world, the harder it is to dig out inconsistencies from it. In this way, blatantly inconsistent worlds have rank 0, while incomplete but consistent impossible worlds and possible worlds have an infinite rank. As for the case of incompleteness, we claimed that we should take each pair of propositions *p* and ~*p*, take the intersection of worlds in them, which results in the singletons of the corresponding lagadonian language sentences, and then for each proposition we should then take the intersection of a world *w* with both singletons, the rank of a world will then be given in term of the number of propositions for which it is not the case that both intersections come out to be the empty set. In a more intuitive way, we can say that the more incomplete world a world is, the less propositions does it represent by containing the relevant lagadonian language sentences. In this context, then, the rank of the blatantly incomplete world will also be 0, whereas maximal worlds (either inconsistent impossible worlds or possible worlds) will have an infinite rank.

In this framework with an extended space of worlds, knowledge is then characterized as truth in all epistemically possible worlds, which may include some inconsistent and/or incomplete worlds. In this way we could provide different truth-conditions for instances such as *Goldbach's Conjecture* and *Goldbach's Conjecture v ~Goldbach's Conjecture* (assuming for the sake of argument that the conjecture is necessarily true). However, epistemic space would be too vast, and the agents would come off as having too little restrictions on what they know given what they know, if only blatantly inconsistent and blatantly incomplete worlds were excluded. For this reason, we then turned to a dynamic approach in which certain closure principles hold, given in terms of what an agent would come to know, if they perform certain actions – that is, we describe a box-like dynamic epistemic operator.

This dynamic epistemic operator, [q], was then defined in terms of what questions an agent considers. The intuitive idea is that considering certain questions guides an agent's inquiry, and leads them to consider a restricted class of worlds. To consider a question is then to have one's epistemic state defined on a restricted space of worlds, the worlds that intuitively "speak to the question". After considering various perspectives on the nature of subject matters, we finally opted for a view according to which subject matters are partitions of the set of worlds speaking to the question, namely a partition where each cell corresponds to an exact truthmaker or falsitymaker for a given statement. We also accepted that the Boolean connectives are transparent in regards to subject matter, and thefore that the subject matter of complex formulae is just the fusion of the subject matters of their atomic constituents. We then maintained that knowledge should be closed under content inclusion, given in terms of truth-conditional inclusion relative to a subject matter and subject matter inclusion. This then allowed us to maintain that knowledge should distribute over conjunction. Finally, it allowed us to keep the notions of subject matter and truth-conditions independent, in the sense of neither determining the other, as Berto et al. (unpublished) argue should be the case.

Objections and replies

While developing the solution for the problem of logical omniscience just presented, we have already considered various objections, providing answers to them as we went on. Still, some objections remain which we now move on to consider.

1. Faulty mereology or misidentification of questions

Earlier we accepted that an agent's knowledge should be closed under content inclusion, that is, that if an agent knows that *p*, if *q*'s truth-conditions include *p*'s truth-conditions and *q*'s subject matter is included in that of *p*, then the agent should likewise know that *q*. As we have noted earlier, however, there might be cases where it is not clear whether an agent's knowledge should be closed in this way. A suggestive example is that knowing that *p* amounts to knowing every proposition where *p* is preceded by an even number of negations. But if the number of negations is very large, for instance one thousand, then it might be that an agent knows that *p* but not that ~...~*p*, where "..." abbreviates nine hundred and ninety eight negation signs. This case is considered as an objection by Hawke et al. (2020, pp. 746-7), as well as an objection against first-degree entailment views in Berto and Jago (2019, p. 117). While both Hawke et al. and Berto and Jago reject a merely metalinguistic explanation of supposedly hyperintensional distinctions induced by agents' epistemic states, like Stalnaker (1984) does, they accept that in certain cases, the best explanation might be metalinguistic - as we claimed earlier, doubts about what proposition a sentence expresses are certainly common, even if they do not seem to account for all the cases that have been considered where an agent apparently has distinct attitudes towards necessarily equivalent propositions. It would seem that in this case it is precisely the complexity of the expression at hand, as in the case explored by Stalnaker (1984, p. 61) that is the reason for doubt. It is not that ~...~*p* has a greater complexity than ~~*p* or *p* in terms of its content, in fact they have the same content in a possible worlds' setting, rather, it is the linguistic expression of the former that is much harder to interpret than the latter's.

This leads us to consider an objection that we have only shortly addressed so far. We have claimed that the subject matters of *p* and of *q* are included in that of *p & q* and in that of *p v q*, which simply correspond to the fusion of the former subject matters. Further, we first claimed that the subject matter of given statements are just certain partitions of sets of worlds that represent that the statement is the case or that it is not the case. But, given our acceptance of Lewis's claim that classes are formed from their proper classes by mereological fusion, then the worlds in the partitions associated with *p* and *q* will not in general correspond to the set of worlds representing *p & q* or *p v q*. For this reason we modified our approach so that the worlds in a partition need not represent a given proposition as being the case, but rather only exact truthmakers/falsitymakers for it. When an agent considers the question of whether *p & q*, for instance, they then can easily rule out worlds not representing *p & q* but which represent *p* and *q* as they are easily recognized as incomplete: they are worlds representing a way for *p & q* to be the case, while not representing it as being the case.

In spite of some of the motivations given earlier, one might still not be satisfied with the present solution. We have previously claimed to be interested in the notion of a question more properly, and only derivatively in the notion of a subject matter, as long as the two are identified. But as we have identified the subject matters of *p & q* and *p v q*, we can no longer identify the subject matters of both with the respective questions *Is p & q the case?* and *Is p v q the case*? for they will correspond to distinct two-celled partitions defined on the same four-celled partition formed by fusing the partition that is the subject matter of *p* with that of *q*. This is to be expected, as we have seen that the identification between a partition containing worlds representing a proposition and its negation with a partition containing worlds representing exact truthmakers/falsitymakers for that proposition only applies in the case of atomic propositions, which only have one exact truthmaker and falsitymaker. Instead of identifying the minimal subject matters of sentences with questions where the answers give simply information on whether a given proposition is the case, the questions must be such that the respective complete answers must give information not on whether they are the case, but on what ways they are/are not the case. We can then think of questions associated with the minimal subject matters of statements to be, for a given associated thin proposition *p*, *How is p the case or not the case?*, where worlds agree on an answer if they contain the same exact truthmaker/falsitymaker for *p*.

Thinking of the question associated with the subject matter of a statement in these terms might not be as intuitive as the approach we started with, so that it might be said that in order to keep upholding the mereology of topics we wished to maintain, that the resulting association between questions and subject matters has become less plausible. Here in defense of this modification I'd like to propose that it is intuitively the case that agents might know what the ways for a proposition to be true/false are, without knowing whether the proposition is true, but on the other hand that if agents know whether a proposition is true, then they must know what its ways of being true/false are. Here note that the requirement is not that the agent should always be able to tell in what way is a given proposition true or false, which as we have seen is violated for instance in some cases where agents know instances of excluded middle but not which of the disjuncts is true. Rather, the requirement is that the agent should know what the ways for it to be true/false are, that is, the agent must know what lagadonian language sentences worlds must contain in order to represent an exact truthmaker or falsitymaker for the proposition. Intuitively, if agents had no grasp on the ways in which a given proposition would be true, then it would seem that they would not be able to tell whether the proposition is true or false, even if given decisive evidence one way or the other. If this is so, then a natural thing to say is that the question of in what ways is a proposition true/false is a part of the question of whether a proposition is the case, but not the other way around. So in this way we can motivate the claim that the questions we have just considered should be taken to be the minimal subject matters of the relevant propositions instead of the candidates previously considered.

Taking the minimal subject matters to be sets of worlds containing exact truthmakers and falsitymakers instead of sets of worlds representing a given proposition or its negation also helps streamline the approach of subject matter inclusion. We've claimed that a subject matter includes another just in case the former is a refinement of the latter. But if subject matters corresponded to *whether* questions, then this would not technically be the case, as in some cases previously considered, the question would not correspond to the four-celled partition itself, but rather, as noted, to a two-cellled partition defined on the four-celled partition. In order to get at a precise theory of subject matter inclusion we'd then have to claim that the minimal question associated with a statement is not the question whose complete answers correspond to the cells of its subject matter partition, but rather corresponding to unions of cells from that partition. The present approach in terms of ways deals, then, with one of the problems facing taking subject matters to be partitions.

2. This conception of content can't accommodate the metalinguistic approach

Having addressed the issue of the present approach supposedly misidentifying questions, let us return to the issue just discussed of cases where an agent seems to know that *p* while failing to know that ~...~*p*. I take it that it is plausible that in such cases, what motivates the differing attitudes is simply a matter of metalinguistic doubt, and therefore that Stalnaker (1984) and Hawke et al. (2020) are right in claiming so. Assuming that they are correct, we should keep holding that the subject matter of a proposition and its negation are identical, and so that the partition of logical space corresponding to **whether *p*** is the same as the partition corresponding to the subject matter of **whether ~...~*p***, so that worlds in both partitions represent *p* and ~*p*, as well as all the other propositions formed out of them by adding an even number of negations. We can start to justify this move by noting that if the metalinguistic approach is right, then to come to know that ~~*p* is no greater achievement than coming to know that *p*, the achievement is to come to know what proposition is expressed by their linguistic expressions, a harder achievement the higher the number of connectives. We can model this extra linguistic complexity in terms of the relevant lagadonian language sentences. To know a given proposition, we have claimed, is for all worlds epistemically possible for an agent to contain the corresponding lagadonian language sentence. If what an agent knows is always relative to a given question, then ~...~*p* will always be known by agents that know that *p*, which seems to accord with the claim that the object of doubt is simply what propositions are expressed by certain sentences. If, on the other hand, agents might have knowledge that is not relative to any subject matter, then they might fail to know that ~...~*p* even if they know that *p*. If we accept this last option, we can still accept a weaker version of the metalinguistic approach: instead of claiming that *p* and ~...~*p* are always known together and that the agent simply fails to know that they know both because of the different linguistic vehicles used to express them, we can claim that there are some contexts in which one but not the other is known, but that the reason for it is simply ignorance that they express the same content relative to any subject matters. On both approaches closure principles for knowledge are given in terms of the dynamic epistemic operator we have been trying to define, but on the first approach these closure principles are then valid for knowledge as such, whereas on the second approach we can only claim that an agent's knowledge is closed whenever they consider certain questions.

The first way of developing the metalinguistic approach relies on the stronger assumption that an agents' epistemic state is always defined relative to a question, so that on this view it's not simply that questions guide inquiry in the sense that by considering questions agents' epistemic states come to be sensitive to the possible complete answers to the question, but further that it is always the case that an agent's epistemic state is sensitive to the complete answers to a given question. In this conception, agents are modelled in a way that resembles to a certain extent the Popperian conception of science. Just as scientists, Popper claims, do not start by making observations which they then hypothesize about, but rather already have certain questions they wish to get answers to when they go out to observe, perhaps all agents are in a similar position in regards to their epistemic lives at large, where it's not that they start by possessing certain items of knowledge, and then coming to have new knowledge based on what questions they consider, but rather that the knowledge they start from is already defined in terms of certain questions they consider. Here we do not take a stance on which of these two approaches has greater plausibility.

As we have suggested earlier, when an agent considers the question of whether *p* is the case, where *p* is a proposition which is not associated with a vague predicate for which one can construct a sorites series with indeterminate borderline cases, then the agent also comes easily to know that *p v ~p*, so that we can explain why it is so easy for agents to come to know *Goldbach's Conjecture v ~Goldbach's Conjecture* while not knowing which of the disjuncts is true. By considering the question of whether *Goldbach's Conjecture* is the case, an agent considers the only two options available: that it is and that it isn't. By doing so, their epistemic state is then focused on the set of worlds speaking to the question, that is, the worlds representing ways for it to be true/false, which is divided into the *Goldbach Conjecture*-worlds and the *~Goldbach's Conjecture*-worlds. This much is given in terms of the dynamic epistemic operator we have been characterizing. In an impossible worlds' setting, however, this partition will not contain only worlds representing *Goldbach's Conjecture v ~Goldbach's Conjecture*, as we have stressed in regards to similar cases. Worlds representing its negation would, as suggested, be ruled out by agents, as contradictions can easily be derived from them, whereas worlds not representing anything in regards to the disjunction would be easily ruled out as incomplete. This entails that coming to know *Goldbach's Conjecture v ~Goldbach's Conjecture* after considering the question of whether *Goldbach's Conjecture* is true has *some* computational cost for agents, even if it is a completely uninformative answer to the question. This corresponds to the step of restricting the set of worlds in the partition that represent ways for the disjunction to be true/false to the set of worlds also representing the disjunction itself. This seems to be the right result: it explains why it is that some agents might suffer from this instance of what Jago (2014, p. 2016) calls an epistemic oversight.

As we've seen earlier, contrary to Jago we do not think of attributions of epistemic oversights to entail that an agent's status as a rational agent has been revoked. Instead, we take it that, after considering the question *Is Goldbach's Conjecture v ~Goldbach's Conjecture the case?* the agent ought to give a positive reply, whenever they have considered the question of whether *Goldbach's Conjecture* is the case, where they recognized such a question as inducing a partition of two exclusive answers, corresponding to the conjecture and its negation. But it might still be the case that even upon considering the question, the agent fails to come to know the disjunction. In spite of such failures, the agent might still be considered to be rational, we've claimed, insofar as they would give the correct answer under optimal conditions, or insofar as they reliably give the correct answer in other relevantly similar cases.

This epistemic oversight seems, however, to be particularly egregious: an agent who knows that *p* and *~p* are the only available answers to *Is p the case?* and afterwards considers **whether *p v ~p*** seems to have an epistemic state first restricted to *p* and *~p* worlds, so that they know either one is the case, which seems to just be what the disjunction expresses. To adopt Fine's terminology, the agent knows that at least one of the verifiers for *p v ~p* holds, so that it seems indeed that they should know that *p v ~p* holds, their ignorance seems to not be about whether it holds but may be rather in terms of in what way they hold (by the actual world being a *p*-world or by the actual world being a *~p* world). This aligns with what has been said earlier in regards to how a world representing a way for a proposition to be the case or not the case but which does not represent it as being the case is easily seen to be incomplete by agents. It is therefore not entirely clear that minimally rational agents might fail in this respect: perhaps in instances where they seemingly fail in this way such agents simply haven't considered the questions they take themselves to have considered, or may have interpreted the linguistic form in which the questions are presented to them in an erroneous way. But even assuming that such epistemic oversights are possible for any agents and therefore that they simply ought to know the disjunction in cases taking this form, we can still explain in the present framework why it seems that they are so egregious.

To uphold the metalinguistic approach in any of the ways earlier described, we had to add the extra requirement that worlds in the partition associated with the subject matter of *p* that represent *p* should also represent that *~~p* and all other such propositions where *p* is preceded by an even number of negations. This seems to be an ad hoc move introduced simply in order for our framework to yield the right results, but we can say more to motivate it, which takes ideas from Yalcin's view of concepts considered earlier. If we think of concepts in the way Yalcin suggests, as we've seen, it would seem that the logical connectives are not plausibly taken to be concepts, as they either seem to behave differently in all possible worlds, or in the same way, so that they would induce no distinctions between possible worlds, or too many. Here we do not take a stance on what the nature of concepts is, but we can adapt the conclusions we have reached in regards to Yalcin's approach if we think of the ways in which the terms of atomic lagadonian language sentences contribute to the distinction induced on the set of worlds which contain it and the lagadonian language sentence formed from the concatenation of it with the negation connective.

Like we've said earlier, if we wish to hold that the connectives contribute in the same way in all worlds to the distinctions induced by the lagadonian language sentences they are parts of, then it would seem that there are no new distinctions between worlds introduced by them, for all the worlds will agree in terms of how the connectives behave in them. On the other hand if we take the connectives to contribute differently across worlds to the distinctions induced by the lagadonian language sentences they are parts of, then it would seem that we still run into the problem of there being too many distinctions that an agent would need to be sensitive to in order to grasp the subject matter of a simple conjunction, for instance. These were the two options identified for Yalcin in the dilemma earlier presented against his position. But in the present context we do not face a dilemma, but rather a trilemma, for there is a new intermediate option between the two options considered earlier of the connectives making no distinctions between worlds, and the connectives making distinctions between each world and every other. Namely, given that complex propositions are attributed a truth-value directly, like atomic propositions, at impossible worlds, there is also the third option of taking the connectives to behave in the usual truth-functional way in possible worlds, but behave differently in impossible worlds, so that there is a middle way between the claim that all worlds are equivalent in regards to the connectives, and the claim that all worlds are distinct from each other in regards to them. But will this third option help with the case just considered of *p* and ~~*p* and others relevantly similar? No, because it will still be the case that all impossible worlds and all possible worlds agree in terms of how the connectives behave. In all possible worlds it will be the case that worlds representing that *p* also represent that ~~*p*, ~...~*p* and all other such propositions. And in the case of impossible worlds, they are all taken to be atomic, so that it would seem that there is no different role that the connectives play in some but not all impossible worlds, and that they are distinguished therefore simply in terms of what worlds contain their respective lagadonian language sentences as members, instead of being distinguished in terms of how the connectives behave across worlds.

To sum up the trilemma, we have that if we want to claim that the connectives contribute to the distinctions between worlds induced by the lagadonian language sentences that contain them, that both the option of taking the connectives to behave in the same way across all worlds and in the same way across all impossible worlds have to be left out for it is in impossible worlds that *p*, ~~*p* and so on can be distinguished, so that if the difference between the corresponding lagadonian language sentences is to contribute to distinctions between worlds representing that ~*p* and ~~~*p*, for instance, then it would seem that the propositions would have to not be evaluated as atomic in impossible worlds, for it would seem that only in that way could then adding negations to a lagadonian language sentence be responsible for distinctions between ~*p* and ~~~*p* worlds. On the other hand, if we allow for the option that the connectives contribute differently across all worlds to the determination of the distinction induced by a lagadonian language sentence, then it would seem that there would be too many distinctions, as for instance worlds representing both *p*, *q* and *p & q* could be distinguished in regards to the subject matter of *p & q* because of whatever else they may represent in terms of the conjunction - for instance if they both represent a further atom *r*, they could be distinguished in terms of one but not the other representing *q & r*, so that the conjunction would thereby behave differently in the two worlds. Intuitively, however, it seems that whether two worlds are equivalent in regards to what *p & q* is about, should not depend on what conjunctions they represent involving content which is not included in that of or related to that of *p & q*.

It seems, then, that the connectives should not be taken to contribute to the distinctions between worlds that certain lagadonian language sentences induce. This gives strength to the claim that Boolean connectives are subject matter transparent: if the connectives do not contribute to distinctions between worlds that can be established by given lagadonian language sentences, then they do not contribute to the determination of the partition induced by such distinctions, that is, they do not contribute to the determination of subject matter. Further, this conclusion motivates a view similar to Hawke's (2017) where the distinctions in the partition are all given in terms of simple tuples of general and individual concepts, namely between the worlds where the individuals fall under the general concept, and those where that condition is not satisfied. In our framework, this corresponds to distinctions in partitions being induced simply by the lagadonian language sentences corresponding to atomic propositions. But finally, and more importantly, this conclusion provides a more solid motivation for the earlier claim that worlds representing that *p* should also represent that ~~*p* and all other propositions which are sets of worlds containing the lagadonian language sentence corresponding to it as part of a lagadonian language sentence formed from it by the addition of an even number of negation connectives. Since the Boolean connectives do not contribute to distinctions between worlds, and therefore do not help determine subject matters, then it would seem that propositions corresponding to lagadonian language sentences differing simply in terms of such connectives should have the same subject matter, that is, that they should correspond to the same partition on worlds. I believe that all these considerations help give some greater plausibility to the claim that worlds in a partition should be closed in this way, which would at first seem to be an ad hoc move.

If in spite of the reasons just provided it still seems plausible to the reader that agents do not always know that ~*...*~*p* when they know that *p* relative to the subject matter of *p*, and so that the metalinguistic approach is not sufficient to account for such cases, perhaps the present framework is also able to accommodate that intuition, with the help of the notions introduced earlier of a partial and full question. As we've claimed, lagadonian language sentences mirror the complexity of the representations we have employed to represent *p* and ~*...*~*p*, so that worlds representing the latter will contain as a member a much more complex lagadonian language sentence. Further, we claimed that for an agent's epistemic state to be defined on a space of worlds where a given proposition is true (that is, for an agent to know a given proposition), then the agent must be able to grasp the lagadonian language proposition that all worlds have in common. For this reason, it might be that agents can only grasp lagadonian language sentences up to a certain level of complexity, so that they might know that *p* but fail to know that ~*...*~*p* simply because they cannot grasp the corresponding lagadonian language sentence. This would then allow us to maintain a still weaker version of the metalinguistic approach: agents might fail to know that ~*...*~*p* in spite of knowing that *p* and this even relative to the question of whether *p*, not because the contents of *p* and ~*...*~*p* are distinct relative to this partition, but rather because the agent might fail to consider the worlds in the partition which contain lagadonian language sentences of a higher syntactic complexity. Following this last strategy, we can identify a further way in which an agent's grasp on a question might be partial: it's not just that the agent might fail to be sensitive to some of the distinctions included in the partition, but further that the cells of the partition that the agent considers might, as it were, contain gaps, gaps which are occupied in the full question by worlds containing certain highly complex lagadonian language sentences that the agent fails to grasp.

Conclusion

Ever since Hintikka proposed to develop an epistemic logic using tools from normal modal logic, the problem of logical omniscience has plagued several modal accounts of knowledge, as well as several epistemic logics constructed in an intensional framework. In response to this challenge, various attempts have been reinterpreted as trying to account not for what an agent explicitly (or even implicitly) knows, but rather for a notion of what an agent is in a position to know. But even if such models are successful in their goals of modelling a notion of what an agent is in a position to know, the notions of implicit and explicit knowledge are central in Epistemology. For this reason, an interest in modelling the less constrained notion of what agents know has been kept alive.

A natural way to tackle various versions of the problem of logical omniscience when trying to model explicit knowledge is to enlarge the space of worlds with what Hintikka called "possible impossible worlds", that is, valuation points that do not respect all the constraints imposed on possible worlds. In such settings an agent's epistemic state might include worlds that represent one but not the other of two necessarily equivalent propositions, so that agents might know one but fail to know the other. Likewise, agents' knowledge will then cease to be closed under logical consequence, so that agents do not come out as knowing all the logical consequences of what they know. But what precisely impossible worlds should represent and what their nature is has been a topic of great discussion, so that our consideration of such approaches to the problem of logical omniscience started by addressing both issues.

While arguing that impossible worlds are entities of a kind with the possible worlds, we took a step back and considered various issues plaguing what we briefly argued, following previous considerations put forth by Jago (2014) and Berto and Jago (2019), to be the best account of worlds: an ersatzist position. Impossible worlds cannot be satisfactorily accommodated in a realist account of worlds, in the sense of Lewis's (1986) modal realism where all worlds are as much a part of reality as the actual world, on the pain of committing the friend of impossible worlds to true contradictions. We moved on therefore to consider proposals that take worlds to be abstract constructions out of actual objects, properties and relations, favouring an approach that takes possible worlds to be maximally consistent sets of lagadonian language sentences, that is, tuples where the objects, properties and relations themselves play the roles of the language's names and predicates. We argued, then, that accepting impossible worlds simply amounted to accepting all other sets of such lagadonian language sentences, and therefore that moving from a possible worlds framework to one containing impossible worlds does not amount to a greater ontological cost. We thereby arrived at a highly anarchic conception of worlds in which what worlds represent is not closed under any weaker-than-classical logic. This allowed us to avoid the trap of accepting that real agents are logically omniscient in respect to one such weaker logic, which we argued was just as unwelcome a version of logical omniscience.

In the process of arguing for our conception of worlds as sets of lagadonian language sentences, we accepted Lewis's criticism of the view that worlds are abstract simples, and that the facts about what worlds represent are themselves primitive and not to be explained in terms of any feature of the worlds. We then saw how Lewis's criticism generalized and seemed to threaten both some versions of ersatzism where worlds were taken to be structured entities but whose simple elements were themselves abstract, as well as the relation of set-membership. We considered several ways of meeting objections and addressing issues in mereology that arise from this difficulty. Resulting from this discussion, we favoured a view of sets in which the members of sets are immediate parts of sets, and sets are fusions of materially-equivalent Fine fusions. This allowed us to hold that the lagadonian language sentences are themselves parts of the worlds that contain them as members, and that therefore on certain views the ultimate constituents of worlds are themselves all concrete, avoiding Lewis's criticism.

Having addressed the aforementioned issues pertaining to the nature of worlds and what they represent, we then moved on to consider how exactly it is that knowledge could be modelled in an impossible worlds' setting, exploring Jago's (2014), Berto and Jago's (2019) and Bjerring and Skipper's (2018) approach. The first two relied on a view of deep epistemic space from which blatantly inconsistent worlds, that is, worlds representing a proposition and its negation, were excluded that we ended up accepting. However, unlike Berto and Jago here it was accepted that worlds may also be excluded from epistemic space on the count of being incomplete, so that blatantly incomplete worlds (namely the empty set of lagadonian language sentences) were also excluded from deep epistemic space. This further way of excluding worlds allowed us to motivate several moves later on in characterizing our own approach to the problem of logical omniscience. Further, Berto and Jago's approach relied on a conception of epistemic space where what is epistemically possible for agents is itself a vague notion. Agreeing with Bjerring and Skipper's inclination to reject this move, we tried to see if there was a way to characterize the knowledge states of minimally rational agents in an impossible worlds setting that did not resort to vagueness. Here we found that the best approach would be to model agents' epistemic states dynamically, so that we would need to characterize a plausible box-like dynamic epistemic operator, that is, we wished to characterize what the agent would come to know after performing a certain action which would change the worlds in an agent's epistemic space. Bjerring and Skipper considered models in which the actions taken by epistemic agents are taking a given number of reasoning steps using certain given rules of inference. But as Berto and Jago rightfully notice, what agents may come to know after some steps of reasoning is too varied, even for very limited sets of rules of inference, so that Bjerring and Skipper fail to characterize a box-like dynamic operator, which is what we were interested in, and instead were only able to characterize a diamond-like operator.

Given how reasoning with certain rules of inference can lead the agent to very different resulting epistemic spaces, we considered instead that the action an agent might take is the action of considering a certain question, following in on the growingly popular tradition of taking agents' mental content to be defined on questions the agents consider. In order to get clearer on how this dynamic operator would be characterized, we then moved on to consider the issue of what the best way is to model the notion of a question and relatedly of a statement's subject matter, topics which have been discussed to a greater extent in recent years after Yablo's (2014) influential theory of aboutness as a separate and irreducible component of meaning.

When considering various ways of understanding the notion of a subject matter, namely Lewis's (1988b), Yablo's (2014), Hawke's (2017), Fine's (2018) and Plebani and Spolaore's (2021), we were met with various suggestions that understanding the notion of a subject matter in an intensional framework does not allow for all kinds of hyperintensional distinctions between propositions which intuitively we would want to make. So for instance in Lewis (1988b) we found the suggestion that adding impossible worlds could potentially help solve the librarian's problem, that is, the problem of untangling such subject matters as **topology** and **algebra**, so that we are able to explain how is it that a librarian is able to store books on different shelves based on the topics they deal with when all propositions in both fields are necessary. Pressured by Fine (2018) (who accepts impossible states from the get go, and we argued has a largely equivalent conception of a space of states to our own), Yablo (2017) admits that his approach merely shows how far it is possible to go in making hyperintensional distinctions before using the "heavy artillery". While Hawke (2017) does not recognize a need for impossible worlds later in Hawke et al. (2020) it is seen that a natural way for his approach to subject matters to untangle the contents of *Goldbach's Conjecture* and of the disjunction whose disjuncts are the conjecture and its negation is to add impossible worlds. Finally, Plebani and Spolaore (2021), like Yablo, admit that their own approach is not able to account for all cases of hyperintensionality, namely not being able to solve the librarian's problem, even if it is able to explain some fallacies of relevance, adding that one might then want to add impossible worlds in order to address remaining issues.

Having examined all of the aforementioned approaches to subject matter, we concluded that in all cases impossible worlds would help precisely in the cases of hyperintensional distinctions we wished to make between necessarily equivalent propositions. We then proceeded to consider two approaches to the problem of logical omniscience that applied the notions of question-sensitivity and subject matter: Yalcin's (2018) and Hawke et al.'s (2020), the first of which accepted a Lewisian conception of subject matters, and therefore suffered from problems previously identified for that conception; whereas Hawke et al. remained uncommitted in regards to what subject matters were, but as we saw, the view that seemed to get closer to the results they wanted to hold was Hawke's (2017). In regards to Yalcin's approach, we noted that it was unable to provide a satisfactory account of why agents may come to know one but not the other of two necessarily equivalent propositions, and further noted that just like Hawke's approach it was not able to distinguish for instance between the contents of *Goldbach's Conjecture* and the disjunction of it and its negation. As it was seen, in order to do so, Yalcin would have to claim that the Boolean connectives were concepts and, following his suggestion, therefore that they induced distinctions between possible worlds, which we considered to be both implausible on its own and to lead to unsatisfactory results if taken seriously.

Contrary to Yalcin's approach, we took Hawke et al.'s (2020) approach to the problem of logical omniscience to get in general things right, and therefore we accepted some of its features, most notably that knowledge distributes under conjunction; and that it is non-monotonic. However, we were not able to fully endorse their approach as it ran into the same problem as Yalcin and Hawke had faced of distinguishing the content of *Goldbach's Conjecture* from the content of the aforementioned disjunction. Further, given that Hawke et al. did not specify what they took to be the nature of topics, the question remained as to whether there were any entities suited to play the role that their theory required of subject matters.

Given our large agreement with Hawke et al.'s approach, our own proposal was to some extent an effort that aimed to provide an account of what entities fulfill the mereology that Hawke et al. imposed on topics, while also validating a form of the principle of closure under conjunction. The resulting view ended up taking elements from all the proposals previously considered of how to understand the notion of a subject matter, where we end up understanding the subject matter of a statement as corresponding to the partition of the set of its exact truthmakers and falsitymakers, so that the cells of the partition corresponded to ways for a given proposition to be true/false. We took these truth/falsitymakers to be the lagadonian language sentences that are members of the worlds, so that our approach started to be very similar to Hawke's. We then moved on to consider various problems in relation to the fact that we were defining subject matters in the context of a very anarchic space of worlds. In order to get at a plausible conception of a question in this setting, we imposed several restrictions on what worlds should be included in a partition: first we noticed that worlds in a partition should be worlds speaking to the question the agent is considering, for otherwise considering a question would not limit epistemic space; second, we considered that by restricting their attention to the worlds in the partition, certain worlds would then be more easily seen to be inconsistent and incomplete, so that they would likewise be easily eliminated from the epistemic space of a minimally rational agent; third we considered that worlds containing lagadonian language sentences with a greater syntactic complexity, but where this complexity is only given in terms of the addition of Boolean connectives should also be contained in the partition of worlds defined in terms of the distinctions induced by the simpler lagadonian language sentences. Besides restricting epistemic space, we also considered such notions as the degree of informativeness an answer carries relative to a question, that limited epistemic agents might only be able to consider partial questions instead of full questions when these entail too many distinctions, or when they contain worlds that have highly complex lagadonian language sentences as members.

Still, by tackling all of these issues, we started to see how it is that subject matters can start to be modeled in a framework with impossible worlds, and how our box-like dynamic epistemic operator meets the various desiderata we identified, guiding the agent's inquiry in the way we required of it. We have arrived then, hopefully, at a plausible solution to the problem of logical omniscience that directly addresses the question of how to model what agents explicitly know. Various other issues were briefly discussed or glossed over, and therefore there is still a lot of work to be done in order to arrive at an answer to the problem of logical omniscience that decides all the details; further, it remains to be seen whether the approach here developed is as successful in dealing with several of the difficulties that have been considered. However, it is our hope that by critically examining existing approaches to the problem of logical omniscience, identifying what seemed to us to be some of its shortcomings, and proposing an alternative that seemingly meets all the desiderata put forth, we have moved closer to a solution to one of the most recalcitrant and long-lasting problems in contemporary formal epistemology.

Bibliography

Armstrong, D. M., *Nominalism and Realism Volume 1: Universals and Scientific Realism*, Cambridge University Press, 1978.

Armstrong, D. M., “Classes are States of Affairs” in Mind, Vol. 100, No. 2, 1991 pp.189-200.

Beall, J. C., Restall, G., *Logical Pluralism*, Oxford University Press, 2006.

Berto, F., “Impossible Worlds and Propositions: Against the Parity Thesis” in The Philosophical Quarterly, Vol. 60, 2010 pp. 471-486

Berto, F., Hawke, P., and Hornischer, L., “Foundations of Two-Component Semantics”, unpublished manuscript, dated March 18, 2021.

Berto, F., and Jago, M., *Impossible Worlds*, Oxford University Press, 2019.

Berto, F., Restall, G., “Negation on the Australian Plan”, in Journal of Philosophical Logic, Vol. 48, 2019 pp. 1119-1144.

Bjerring, J.C., “Impossible worlds and logical omniscience: an impossibility result” in Synthese, 2012.

Bjerring, J.C. and Skipper, M., "A Dynamic Solution to the Problem of Logical Omniscience" in Journal of Philosophical Logic, 2018.

Bøhn, E. D., “Commentary on “Parts of Classes”” in Humana.Mente Journal of Philosophical Studies, Vol. 19, 2011 pp.151–158.

Bricker, P., “Island Universes and the Analysis of Modality” in *Reality and Humean Supervinience. Essays on the Philosophy of David Lewis*, (ed. and int.) Gerhard Preyer and Frank Siebert, 2001 pp. 27-55

Bunt, H. C., *Mass Terms and model-theoretic semantics*, Cambridge University Press, 1985

Button, T., and Walsh, S., *Philosophy and Model Theory*, Oxford University Press, 2018.

Caplan, B., Tillman, C. and Reeder, P., “Parts of Singletons” in The Journal of Philosophy, Vol. 107, No. 10, 2010 pp. 501-533.

Carlson, K., “Clefting, Parallelism, and Focus in Ellipsis Sentences”, in (eds.) L. Frazier and E. Gibson, *Explicit and Implicit Prosody in Sentence Processing*, Springer, 2015 pp. 63-83.

Chalmers, D., “The nature of epistemic space” in A. Egan and B. Weatherson (eds), *Epistemic Modality*, Oxford University Press, 2010 pp. 60–107.

Cotnoir, A. J., and Bacon, A., “Non-wellfounded Mereology” in The Review of Symbolic Logic, Vol. 5, No. 2, 2012 pp. 187-204.

Cotnoir, A. J., “Is Weak Supplementation analytic?” in Synthese, 2018.

Cross, C., and Roelofsen, F., “Questions” in Stanford Encyclopedia of Philosophy, 2018 https://plato.stanford.edu/entries/questions/ Last consulted on 06/08/2021.

Denby, D. A., “In defence of Magical Ersatzism”, The Philosophical Quarterly, Vol. 56, No. 223, 2006 pp.161-174

Fine, K., “First-Order Modal Theories III: Facts”, in Synthese, Vol. 53, No. 1, 1982, pp. 43–122.

Fine, K., “Essence and Modality: The Second Philosophical Perspectives Lecture”, in Philosophical Perspectives, Vol. 8, Logic and Language, 1994, pp. 1-16.

Fine, K., “Things and their Parts” in Midwest Studies in Philosophy, Vol. 23, No. 1, 1999 pp. 61-74.

Fine, K., “The problem of possibilia” in *The Oxford Handbook of Metaphysics*, (ed.) Michael J. Loux and Dean W. Zimmerman, Oxford University Press, 2005, pp. 161-179.

Fine, K., “Towards a Theory of Part” in The Journal of Philosophy, Vol. 107, No. 11, 2010 pp. 559-589.

Fine, K., “Yablo on Subject-matter”, in Philosophical Studies, Vol. 177, No. 1, 2018.

Fine, K., and Jago, M., “Logic for exact entailment” in Review of Symbolic Logic, Vol. 12, No. 3, 2019 pp. 536-556.

Fine, K., “Constructing the Impossible”, in (eds.) Lee Waters and John Hawthorne, *Conditionals, Paradox, and Probability: Themes from the Philosophy of Dorothy Edington*, Oxford University Press, 2021.

Fisher, A. R. J., “On Lewis against magic: a study of method in metaphysics” in Synthese, 2015

Frege, G., “The Thought”, in Mind, Vol. 65, No. 259, 1956, pp. 289-311.

Goldblatt, R., *Lectures on the Hyperreals: An Introduction to Nonstandard Analysis*, Springer, 1998.

Hawke, P., “Theories of Aboutness”, in Australasian Journal of Philosophy, Vol. 96, No. 4, 2017, pp. 697-723.

Hawke, P., Özgün, A., and Berto, F., “The Fundamental Problem of Logical Omniscience”, in Journal of Philosophical Logic, Vol. 49, 2020, pp. 727–766.

Hintikka, J., *Knowledge and Belief: An Introduction to the Logic of the Two Notions*, Cornell University Press, 1962.

Hintikka, J., “Impossible possible worlds vindicated”, in Journal of Philosophical Logic, Vol. 4,1975 pp. 475-484.

Hoek, D., “Minimal Rationality and the Web of Questions”, forthcoming in Kindermann, Van Elswyk and Egan (eds.), *Unstructured Content*, Oxford University Press.

Jacinto, B., "Possible Worlds" in Online Companion to Problems in Analytic Philosophy, Centro de Filosofia da Universidade de Lisboa, 2013.

Jago, M., *The Impossible: An Essay on Hyperintensionality*, Oxford University Press, 2014.

Jago, M., "Hyperintensional Propositions", in Synthese, No.192, 2015, pp.585–601.

Johnston, M., "Hylomorphism", in Journal of Philosophy, Vol. 103, No. 12, 2006, pp. 653–698.

Jubien, M., "Could this be Magic?", The Philosophical Review, Vol. 100, No. 2, 1991 pp. 249-267.

Kaplan, D., "Dthat", in *Syntax and Semantics*, (ed.) P. Cole, Vol. 9, 1978, pp. 221–53.

King, J., "Structured propositions and complex predicates", in Noûs, Vol. 29, pp. 516–35.

Kratzer, A., "Situations in Natural Language Semantics" in Stanford Encyclopedia of Philosophy, 2019.

Künne, W., *Conceptions of Truth*, Oxford University Press, 2003.

Lewis, D., *Counterfactuals*, Blackwell, 1973.

Lewis, D., "Logic for Equivocators", in Noûs, Vol. 16, No. 3, 1982, pp. 431-441.

Lewis, D., *On the Plurality of Worlds*, Blackwell, 1986.

Lewis, D., "Statements partly about observation", Philosophical Papers, Vol. 17, No. 1, 1988a, pp. 1-31.

Lewis, D., "Relevant Implication", in Theoria, Vol. 54, No. 3, 1988b, pp. 161-174.

Lewis, D., *Parts of Classes*, Blackwell, 1991.

Lewis, D., "Elusive Knowledge" in Australasian Journal of Philosophy, Vol. 74, No. 4, 1996 pp. 549-567.

Lewis, D., "Letters to Priest and Beall" in B. Armour-Garb, J. Beall and G. Priest (eds.), *The Law of Non-Contradiction: New Philosophical Essays*, Oxford University Press, 2004 pp. 176–7.

MacFarlane, "Frege, Kant, and the Logic in Logicism", in The Philosophical Review, Vol. 111, No. 1, 2002 pp. 25-65.

Menzel, C., "Problems with the Actualist Accounts" in Stanford Encyclopedia of Philosophy, 2014. https://plato.stanford.edu/entries/actualism/actualist-problems.html. Last accessed on 31/05/2021.

Nolan, D., "Impossible Worlds: A Modest Approach", Notre Dame Journal of Formal Logic, Vol. 38, No. 4, 1997, pp. 535-572.

Orilia, F., and Paoletti, M. P., "Properties" in Stanford Encyclopedia of Philosophy, 2020 https://plato.stanford.edu/entries/properties/#SparAbunConc Last consulted on: 22/07/2021.

Paseau, A., “Motivating Reductionism about Sets” in Australasian Journal of Philosophy, Vol.86, No.2, 2008 pp.295-307.

Plantinga, A., *The Nature of Necessity*, Oxford University Press, 1974.

Plato, *Euthyphro. Apology. Crito. Phaedo. Phaedrus*, (int.) W. R. M. Lamb, (trans.) Harold North Fowler, Harvard University Press, 2015

Plebani, M., and Spolaore, G., “Subject Matter: A Modest Proposal”, in The Philosophical Quarterly, Vol. 71, No. 3, 2021, pp. 605-622.

Priest, G., *Towards non-being: The Logic and Metaphysics of Intentionality*, Oxford University Press, 2005.

Priest, G., *An Introduction to Non-Classical Logic*, Cambridge University Press, 2012.

Priest, G., “Mission Impossible”, forthcoming in *Saul Kripke on Modal Logic*, (eds.) Yale Weiss and Romina Padró, Springer.

Quine, W. V., “Three grades of modal involvement” in *The Ways of Paradox and other essays*, Harvard University Press, 1976 (first published in 1953).

Sider, T., *The Tools of Metaphysics and the Metaphysics of Science*, Oxford University Press, 2020.

Stalnaker, R., “Possible Worlds”, in Noûs, Vol. 10, No. 1, 1976, pp. 65-75.

Stalnaker, R., *Inquiry*, MIT Press, 1984.

Stalnaker, R., “Impossibilities” in Philosophical Topics, Vol. 24, No. 1, 1996, pp. 193-204.

Stalnaker, R., *Mere Possibilities. Metaphysical Foundations of Modal Semantics*, Princeton University Press, 2012.

Textor, M., “States of Affairs” in Stanford Encyclopedia of Philosophy, 2021. https://plato.stanford.edu/entries/states-of-affairs/#StatAffaComp . Last accessed on 31/05/2021.

Thomson, J. J., “The statue and the clay”, in Noûs, Vol. 32, No. 2, 1998, pp. 149–173.

Van Inwagen, P., “Two Concepts of Possible Worlds”, Midwest Studies in Philosophy No. 11, 1986 pp. 185-213.

Varzi, A., “Mereology” in Stanford Encyclopedia of Philosophy, 2016 https://plato.stanford.edu/entries/mereology/ Last consulted in: 22/07/2021.

Williamson, T., *Knowledge and its Limits*, Oxford University Press, 2000.

Williamson, T., *The Philosophy of Philosophy*, Oxford University Press, 2007.

Wittgenstein, L., *Tractatus Logico-Philosophicus*, Routledge and Kegan Paul, 1921.

Yablo, S., *Aboutness*, Princeton University Press, 2014.

Yablo, S., “Reply to Fine on *Aboutness*”, in Philosophical Studies, Vol. 175, 2017 pp. 1495-1512.

Yagisawa, T., “Worlds and Individuals, Possible and Otherwise”, Oxford University Press, 2010.

Yalcin, S., “Belief as Question-Sensitive”, in Philosophy and Phenomenological Research, Vol. 97, No. 1, 2018 pp. 23-47.

Yalcin, S., “Fragmented but rational”, manuscript of 8th of July 2020, forthcoming in *The Fragmented Mind*, (eds.) Cristina Borgoni, Dirk Kindermann, and Andrea Onofri, Oxford University Press, 2021.

Zaragoza, K., “Bring Back the Magic”, Pacific Philosophical Quarterly, No. 88, 2007, pp.391–402.