Università degli Studi di Padova

Head Office: Università degli Studi di Padova

Department of Philosophy, Sociology, Education and Applied Psychology (FISPPA)

Ph.D. COURSE IN PHILOSOPHY
SERIES XXXIV

Gluons, rejection, and other dialetheic issues: new perspectives

Coordinator: Prof. Luca Basso
Supervisor: Prof. Massimiliano Carrara

Ph.D. student: Filippo Mancini

"The test of a first-rate intelligence is the ability to hold two opposed ideas in the mind at the same time, and still retain the ability to function"

F. Scott Fitzgerald, *The Crack-Up*

Abstract

Dialetheism is the view that there are true contradictions – i.e., both true and false sentences. It is a controversial thesis, and represents a genuine metaphysical revolution. Also, it might be very fruitful from several points of view. It seems to bring some remarkable advantages. Among them, dialetheism can be considered a 'solution' to the whole family of the paradoxes of self-reference – such as the liar and Russell's paradoxes. Or, it might address some actual inconsistent *phenomena*, such as the transitions states and the topological boundaries. Again, it allows the development of an entirely new approach to mathematics, known as inconsistent mathematics, of which very little has been explored by logicians and mathematicians until now. However, there are also significant disadvantages, and the possible success of dialetheism strictly relies on it being cost-effective – i.e. rationally preferable – compared to its competing theories. Thus, this work aims to help the assessment of dialetheism, and to give a contribution to tipping the scales in favor or against such a view.

The main problem with dialetheism is known as the *exclusion problem*. In short, the formal dialetheic semantics lacks any exclusive connective, but it is acknowledged that every logic aspiring to be correct must include at least one. Prof. Graham Priest, one of the fathers of dialetheism, has proposed to recover exclusivity at the pragmatic level – in terms of the propositional attitudes of belief and rejection (and their related speech acts of assertion and denial). He claims that these pairs of notions are exclusive (i.e., incompatible) and that do not allow new paradoxes to emerge, being they uniquely formalized as force operators. However, I think this is too far. For in chapter 2 I prove and discuss two paradoxes – the denial and the rejection paradoxes – that originate from the notions of rejection and denial embedded as predicates. To prove them, I use the formal logic DLEAC, which faithfully mimics Priest's pragmatics. Therefore, these paradoxes undermine his solution to the exclusion problem and consequently weaken the alleged supremacy of dialetheism.

Another good point of dialetheism is that it enables the development of new metaphysical theories to face some long-standing metaphysical problems. This is the case of gluon theory, that Prof. Priest devised to solve the unfamous problem known as *the one and the many problem*. Such a theory is the result of combining some more basic theories, namely: dialetheism, a non-well-founded mereology and modal meinongianism. It seems to solve the problem for which it was conceived – by breaking the infinite vicious Bradley regress. However, in chapter 3 I show that from this theory we can prove a very problematic theorem, which informally states the inconsistency of every object – i.e., that every object has at least one contradictory property. Then, this result put pressure on the validity of gluon theory, as well as on dialetheism, as the most directly involved of its components.

Finally, in chapter 4 I sketch and discuss a possible outline for an inconsistent megethology. This can be thought of as an alternative foundation to inconsistent mathematics: by combining an inconsistent mereology and plural quantification, it aims at expressing naive set theory as D. Lewis' megethology expresses ZFC set theory.

Acknowledgements

I would like to express my gratitude to my supervisor, Massimiliano Carrara, who guided me throughout my doctoral research. I would also like to thank Graham Priest, Naoya Fujikawa, Filippo Casati, and Jeroen Smid for their cooperation and very helpful comments and insights.

Contents

List of Figures

List of Tables

To the peace of all sentient beings

Chapter 1

Dialetheism: a (not so) short introduction

1.1 Preliminary remarks

There is a very important matter we need to begin with: belief-revision. This is the process through which an ideal rational agent revises her own beliefs to get an ever-improving understanding of the world, i.e. a better representation of it. How does this process (should) work? There is a well-known formal account that gives a model of it: the AGM theory.[1] Although such an account is not unproblematic and despite being there different proposals on the market, it is the currently dominant view. Thus, we will briefly discuss it since it will allow us to put the main topic of this chapter – i.e. dialetheism – into perspective. More precisely, our focus here is theory choice, which can be thought of as an application of belief-revision on a set of competing theories.[2] We are interested in the following question: given two competing theories both aiming at accounting for the same collection of *phenomena*, which one should we prefer/believe? Easy: the better theory! But of course, this is just a vacuous answer since it only postpones the question. For how to know which is the better theory? As a matter of fact, what we need is a set of *criteria* that enable us to evaluate the competing theories. We need to "mark" every theory so as to arrange all of them into a strict descending order: the first theory of the sequence is the one we want to endorse/believe. That is all well and good, but the task is far from easy.

In the AGM theory, the beliefs are represented by sentences in a formal language, $\mathcal{L}$. In particular, the beliefs held by a rational agent are represented by K, which is a set of sentences of $\mathcal{L}$. It is usually assumed that K is closed under logical consequence.[3] This move is quite unrealistic since it makes the rational agent "logically omniscient". However, it is very useful and it can be easily handled by considering K as the set of sentences that the agent is *committed* to believing. Now, given that the

[1] The theory is named after its fathers: Carlos Alchourrón, Peter Gärdenfors, and David Makinson. See Hansson (2017) for an introduction.

[2] A (formal) theory is a logically closed set of (formal) sentences.

[3] The kind of logical consequence relation (i.e. logic) to be used is a free parameter, except for some possible changes the theory may require depending on the logic we opt for. The standard choice is some version of classical logic.

aim is to model the process of belief-revision, we need to include some operations on K which represent the belief changes. In the AGM account, there are three: expansion ($+$), contraction ($-$), and revision ($*$). Expansion models the addition of a belief, say α, when nothing is removed: K is replaced by $K + \alpha$, that is the smallest logically closed set that contains both K and α. Thus, $K + \alpha = \{\beta : K \cup \{\alpha\} \vdash \beta\}$, where $\vdash$ denotes the (selected) consequence relation. Contraction models the removal of a belief. This is not just to delete α from K. Since the result must be logically closed, we may have to delete other things as well. From K we get $K - \alpha$, that is a set such that $K - \alpha \subseteq K$ and that $\alpha \notin K - \alpha$, but this change can be accomplished in different ways – i.e. there are many sets $K - \alpha$ satisfying these conditions. The AGM account does not give an explicit definition of contraction but gives a set of axioms that $K - \alpha$ must satisfy. These so-called basic AGM postulates are the following:

(Closure$_-$) $K - \alpha = \{\beta : K - \alpha \vdash \beta\}$

The result of the contraction operation is logically closed.

(Success$_-$) If $\alpha \notin \{\beta : \vdash \beta\}$ then $\alpha \notin K - \alpha$

If α is not a logical truth, then it is not implied by the sentences of $K - \alpha$.

(Inclusion$_-$) $K - \alpha \subseteq K$

The contracted set is a subset of the non-contracted one.

(Vacuity$_-$) If $\alpha \notin \{\beta : K \vdash \beta\}$ then $K - \alpha = K$

If the sentence to be contracted is not included in the original belief set, then contraction by that sentence involves no change at all.

(Extensionality$_-$) If $\alpha \leftrightarrow \gamma \in \{\beta : \vdash \beta\}$ then $K - \alpha = K - \gamma$

The result of contracting by logically equivalent sentences is the same.

(Recovery$_-$) $K \subseteq (K - \alpha) + \alpha$

So much is retained after α has been removed that everything will be recovered by reinclusion (through expansion) of α. This guarantees that belief contraction is *minimal* in the sense of leading to the loss of as few previous beliefs as possible.

(Conjunctive inclusion$_-$) If $\alpha \notin K - (\alpha \wedge \beta)$ then $K - (\alpha \wedge \beta) \subseteq K - \alpha$

Everything that is retained in $K - (\alpha \wedge \beta)$ is also retained in $K - \alpha$.

(Conjunctive overlap$_{-}$) $(K - \alpha) \cap (K - \beta) \subseteq K - (\alpha \wedge \beta)$

Whatever can withstand both contraction by α and contraction by β, can also withstand contraction by $(\alpha \wedge \beta)$.

Revision models the addition of a belief to K when other sentences have to be removed to ensure that the resulting set of beliefs, $K * \alpha$, is consistent. As for contraction, also revision has been axiomatically characterized. Thus, we have the following postulates:

(Closure$_{*}$) $K * \alpha = \{\beta : K * \alpha \vdash \beta\}$

The result of the revision operation is logically closed.

(Success$_{*}$) $\alpha \in K * \alpha$

The agent is committed to believing α after the revision process by α.

(Inclusion$_{*}$) $K * \alpha \subseteq K + \alpha$

The revised set of beliefs is a subset of the expanded set (by the same sentence).

(Vacuity$_{*}$) If $\neg\alpha \notin K$ then $K * \alpha = K + \alpha$

If the agent does not believe $\neg\alpha$, then revision boils down to expansion.

(Consistency$_{*}$) If α is consistent, then $K * \alpha$ is consistent

The outcome of revising by a consistent sentence is a consistent set of beliefs (provided K is consistent).

(Extensionality$_{*}$) If $\alpha \leftrightarrow \gamma \in \{\beta : \vdash \beta\}$ then $K * \alpha = K * \gamma$

The result of revising by logically equivalent sentences is the same.

(Superexpansion$_{*}$) $K * (\alpha \wedge \beta) \subseteq (K * \alpha) + \beta$

The set of beliefs after revising by a conjunction is a subset of the expansion by one of the conjuncts of the original set revised by the other conjunct.

(Subexpansion$_{*}$) If $\neg\gamma \notin \{\beta : K * \alpha \vdash \beta\}$ then $(K * \alpha) + \gamma \subseteq K * (\alpha \wedge \gamma)$

If after the revision by α the agent is not committed to $\neg\gamma$, the expansion of such a set of beliefs by γ is a subset of the revision of the original set, K, by the conjunction between α and γ.

Levi (1978) has shown that if we define revision in terms of expansion and contraction as follows:

$$\text{(Levi Identity)} \qquad K * \alpha = (K - \neg\alpha) + \alpha,$$

then the AGM conditions for $+$ and $-$ entail those for $*$.

Now, it is important to note that the AGM account describes the process of belief-revision, but is silent about when it is rational to perform such a process. Let us consider the following situation. I am a rational agent who holds a set of beliefs, K. Among them, some form a theory, T, which describes a certain class of *phenomena*. Therefore, $T \subset K$. Then, I find out about a new theory, T', which describes the very same class of *phenomena* of T, and such that $T \neq T'$. Should I revise my beliefs in favor of T' or retain T and discard the new theory? The AGM account does not give us any guidance about that. However, as we suggested before, to answer the question, we need to identify the features of a set of beliefs that may speak in favor or against its rational acceptability, and rate the competing theories based on them. Arguably, the following list includes the most relevant of such *criteria*: the explanatory power, the empirical adequacy, the way it coheres with the extant net of knowledge, the predictive power, the paucity of *ad hoc* hypothesis, the elegance and simplicity. Unfortunately, there are at least three difficulties behind the corner: (i) we do not know whether or not the list of *criteria* we have just sketched is exhaustive, (ii) it is far from easy to quantitatively evaluate such *criteria*, and (iii) the evaluation appears to be rather vague and subjective. In short: there seems to be no algorithmic procedure to declare the winning theory. Nevertheless, we can still try to formally – and roughly – capture the process of theory choice. For that, we will follow Priest (2006a, §8.6).

Let K our set of beliefs and α the new incoming information. Then, what is the new rational set of beliefs? Well, it depends. As we have seen, there are many actions we can take: for example, we can expand K to get $K + \alpha$, we can revise it to get $K * \alpha$, or we can reject α. Further, there are different revision operations, depending on how we define contraction. Moreover, another possibility – which is not modeled by the AGM account – is conceptual innovation. This move requires changing the language $\mathcal{L}$ to improve its expressive power. Let us call the collection of all these possibilities $K^\alpha = \{K_1, K_2, ..., K_m\}$. Then, take $C = \{c_1, c_2, ..., c_n\}$ to be the set of *criteria* that can be used to evaluate a set of beliefs. Let assume that for every $K_j \in K^\alpha$, where $j = 1, ..., m$, and for every $c_i \in C$, such that $i = 1, ..., n$, there is an interval $\mu_{c_i}(K_j) = [\mu^-_{c_i}, \mu^+_{c_i}]$ which represents the vague score we assign to K_j with respect to that particular *criterion*, c_i. $\mu^-_{c_i}$ is the lower bound of the interval, i.e. the minimal score we are willing to assign, whereas $\mu^+_{c_i}$ is the upper bound, i.e. the maximal score. Now, to get the total score of K_j we can just calculate the weighted average of $\mu_{c_i}(K_j)$. This will be an interval,

$\rho(K_j) = [\rho^-(K_j), \rho^+(K_j)]$, where:

$$\rho^-(K_j) = \sum_{1 \le i \le n} w_i \mu^-_{c_i}(K_j)$$

$$\rho^+(K_j) = \sum_{1 \le i \le n} w_i \mu^+_{c_i}(K_j)$$

Here, the weights $w_i \in [0,1]$ reflect the relative importance of each *criterion*.[4] Next, we define an overall ranking, $\sqsubset$, on K^α. As Priest (2006a, p. 137) puts it: «[o]ne belief-set, K_1, is rationally preferable to another, K_2, if it is clearly better, that is, if any value that K_1 can have is better than any value K_2 can have» – i.e. if the two intervals do not overlap each other. Formally:

$$K_2 \sqsubset K_1 \Leftrightarrow \rho^-(K_1) > \rho^+(K_2)$$

$\sqsubset$ is a partial order but is not, in general, a total order. Thus, «it may be that the rationality ordering, $\sqsubset$, provides a clear judgment concerning two theories sometimes, but not others. Rational disagreement is possible» (Priest, 2006a, p. 138). Finally, we define the rational set of beliefs – which corresponds to the best choice we can make –, $K_{Rat} \in K^\alpha$, (one of) the maximal set(s) of the ordering. If there is more than one, the situation is non-deterministic.

Of course, the theory we have just sketched is just that: a theory. Some significant complications arise from the fact that we do not have a complete set of *criteria*, C, and that the weights w_i may vary with time and contexts. However, it provides a good enough idea about how rational belief-revision works. Moreover, it is very general: it can be applied to whatever subject matter we are interested in. For example, it can model the belief-revision in physics, as well as in philosophy or logic. Arguably, the only difference is the set of *criteria* we have to evaluate. For example, empirical adequacy is likely (one of) the most important features that a physical theory must satisfy. However, those who claim that logic and philosophy are *a priori* disciplines would agree that this is hardly a property we can examine for these subjects.

Now, let us call consistency into play. A theory is consistent if it does not entail any contradiction. It is a standard practice to consider consistency as a necessary feature for every theory which aims at being rationally eligible. According to this view, inconsistency is an unquestionable mark of irrationality. Therefore, every inconsistent theory must be rejected. But this is too fast, if not downright wrong. In light of the model of belief-revision we have just discussed, we can think of consistency as one of the many other evaluation *criteria* we have already introduced. As we

[4] In this model, we assume that the *criterion*-weights are real numbers. However, they are not uniquely and precisely determined in general. Therefore, a more realistic model can be obtained by assuming that even the weights are intervals.

will discuss in §1.2, it comes by degrees, and even considering it a positive feature, not necessarily the most rational choice will go in the direction of a fully consistent theory. Think about a theory that scores very little in consistency but very much in all the other *criteria*. This could be better than another theory which scores very much in consistency but very little in the remaining features. Thus, it is because rationality is a matter of pondering multiple *criteria* that inconsistency may also prove to be the best choice.[5] Of course, it is important to note that if we allow for inconsistency there are some changes we need to make to the AGM theory. For such a kind of revision is not properly captured by the AGM axioms. Consider for example (Consistency$_*$): it must fail because the revised set of beliefs $K * \alpha$ may be inconsistent even if α is consistent. But we do not need to go into the details here.[6] The moral should be clear: rationality and inconsistency can be compatible. Therefore, we may consider going for inconsistency. We just need to face the following question: from a rational point of view, does it worth it?

1.2 Dialetheism and its logic

Dialetheism – sometimes also called strong paraconsistency – is the metaphysical view that there are true contradictions.[7] A contradiction is a sentence[8] of the form: α and it is not the case that α. Using the ordinary symbols for conjunction ($\wedge$) and negation ($\neg$), the form of a contradiction is $\alpha \wedge \neg\alpha$. Inspired by the Janus-headed nature of contradictions,[9] Graham Priest and Richard Routley, the fathers of dialetheism, coined the neologism «*dialetheia*» for «true contradiction». Thus, for a dialetheist there are dialetheias. If we assume (the standard clauses) that a statement is false if and only if (from now on, iff) its negation is true, and that a conjunction is true iff both of its conjuncts are true, another way to express dialetheism is to say that there are both true and false sentences. This work aims to help to assess dialetheism, that is to make a contribution to tipping the scales in favor or against such a view.

Dialetheism requires a very peculiar logical setting – i.e. a dialetheic logic, where logic should be taken in a broad sense, including a formal logical theory, a theory of truth and pragmatics. The reason why some logical changes are needed is that

[5]Though it depends on the level of inconsistency. For example, as we will discuss in the next section, triviality is an extreme form of inconsistency we ought to reject.

[6]See Tanaka (2005) for a supplementary analysis.

[7]It is important to notice that, even if dialetheism is a unique and precise thesis, two dialetheists may disagree about some more specific topics. For instance, the dialetheias hold by the first dialetheist may conflict with the dialetheias hold by the second one – i.e., they disagree on which contradictions are true in the actual world. Or, they can accept different paraconsistent logics. Therefore, even if the core idea – dialetheism – is maintained, there is room for (sometimes considerable) disagreement. Arguably, the two most important dialetheic views are those developed by Graham Priest and Jeffrey C. Beall. Because of its role and the deep impact it has had in the debate so far, I will almost exclusively focus on Priest's dialetheism.

[8]Or statement, or proposition, or any truth-bearer you prefer. It does not make a significant difference for our purpose.

[9]Wittgenstein (1956, pt IV, sect. 59)

you cannot be both a dialetheist and a classical logician at the same time, on pain of triviality. Trivialism is the view that every sentence is true. Most philosophers reject it,[10] Priest included: «belief in [trivialism], though, would appear to be grounds for certifiable insanity» (Priest, 1999, p. 443). Many counter-arguments can be given against trivialism.[11] To mention just one:

> One cannot intend to act in such a way as to bring about some state of affairs, *s*, if one believes *s* already to hold. Conversely, if one acts with the purpose of bringing *s* about, one cannot believe that *s* already obtains. Hence, if one believes that everything is true, one cannot act purposefully.
>
> Priest (2006a, p.69)

Classical logic validates an inference schema called *ex contradictione quodlibet*, or the principle of Explosion, according to which any contradiction implies anything: for every α and β, $\{\alpha, \neg\alpha\} \models \beta$. Because of that, the logical consequence relation ($\models$) is said to be explosive. But now, if we allow for some contradiction to be true, we make every sentence true through Explosion, i.e., we get triviality. Thus, to avoid such an undesirable consequence, we need to make some adjustments.

To introduce dialetheic logic, we will closely follow Priest (2006b, Part II), which has become the canonical manifesto of dialetheism.

1.2.1 The teleological theory of truth

The first point to be discussed is the theory of truth, together with some related notions. As Priest (2006b, p. 53) makes clear, «[d]ialetheism [...] does not commit one *per se* to any particular account of truth». In addition, Priest (2006a, §2) claims that the main traditional accounts of truth[12] are all quite compatible with the existence of true contradictions. However, there are some issues about truth that require to be addressed, the most important of which is what is to say of something that is true.

A preliminary remark is that Priest takes truth to be attributed to sentences (without indexicals). He avoids talking about other truth-bearers such as statements, propositions, beliefs, and other cognitive entities. This said, Priest takes truth to satisfy the T-scheme:[13]

$$T\underline{\alpha} \leftrightarrow \alpha \tag{1.1}$$

[10] Though, there have been some attempts to defend it. See for example Kabay (2008).

[11] See for example Priest (2006a, §§3.5-9). Also, note that any argument against dialetheism is *ipso facto* an argument against trivialism. For believing trivialism implies believing dialetheism (even if the converse does not hold) and, *via modus tollens*, rejecting dialetheism entails rejecting trivialism as well.

[12] The correspondence, coherence, pragmatist, deflationist, semantic and teleological theories of truth.

[13] Such a scheme is well known in philosophy and likely does not need any presentation. The reader who is not familiar with it can see Hodges (2018) or Tarski (1956).

Here, T is the truth predicate, $\underline{\alpha}$ is the name of α[14] and what exactly $\leftrightarrow$ is should become clear later on. There are some arguments[15] supporting this claim, one of which is that the T-scheme properly captures the "disquotational" features of truth, which are crucial for this notion. But can we say that the T-scheme produces an adequate characterisation of truth *simpliciter*? In other words: if some predicate satisfies the T-scheme, is it *ipso facto* truth? If the connectives in the T-scheme were extensional, then the scheme would characterise truth extensionally and, in general, this is not enough. For there are different notions with the same extension, i.e., having the same extension does not guarantee the sameness of notions. However, as it will be discussed in §1.2.2, the (bi)conditional occurring in the scheme is stronger than a material one. Priest opts for a non-detachable material conditional ($\supset$),[16] but he does want detachability in this case. Therefore, the conditional he chooses is the stronger genuine conditional ($\rightarrow$) he defines as an intensional connective. Nevertheless, despite the genuine conditional being intensional, the T-scheme on its own is not enough to give an adequate characterisation of truth: «[i]t is the use to which the truth predicate is put, and in particular its connection with the things that speakers wish to or are prepared to assert, that completes its characterisation» (Priest, 2006b, p. 62). Thus, according to Priest the truth is the *telos* of the assertoric speech act, i.e., the aim of asserting. In virtue of that, we can refer to Priest's account as a *teleological* account of truth. Of course, this view requires a good insight of asserting. In particular, a suitable notion of assertion must not be based on truth, on pain of circularity. We will discuss it in §1.2.3, where we will introduce Priest's pragmatics.

This is a convenient time to face some other issues, precisely: the notions of falsity and untruth, and the alleged existence of valueless sentences. Priest takes falsity (F) to be the truth of negation. That is, for every sentence α:

$$F\underline{\alpha} \leftrightarrow T\underline{\neg\alpha} \tag{1.2}$$

Falsity is then defined in terms of truth and negation, and this move is legitimate as long as we do not define negation in terms of falsity – which is the orthodox strategy. Priest does not offer an independent characterization of negation, but this is not something we should be worried about. For something similar would also happen for the classical logician. She considers truth and falsity independently definable notions and defines negation based on them: a sentential truth-value flipping function. But then she struggles in grounding truth and falsity independently. This makes Priest conclude she is not in a better position than him, which would be a good enough reason to move forward.

[14] I will use underlining as a quotation device throughout the whole thesis.

[15] See Priest (2006b, §1.4 and §§4.2-3).

[16] A non-detachable conditional is a conditional that invalidates *modus ponens*.

As we said, dialetheism endorses truth-value *gluts*, i.e. both true and false. Does it also allow for truth-value *gaps*, i.e. neither true nor false? According to Priest, the answer is in the negative. Cases of reference failure, paradoxical sentences, and all the other situations usually considered as examples of valueless sentences are, in his view, to be rejected. A brief justification can be given by quoting his own words: «[s]uppose that α is a sentence, and suppose that there is nothing in the world in virtue of which α is true – no fact, no proof, no experimental test. Then this is the Fact in virtue of which $\neg\alpha$ is true» (Priest, 2006b, p.64). Since if $\neg\alpha$ is true then α is false, truth and falsity are exhaustive: the Law of Excluded Middle (LEM) – i.e. $\models \alpha \vee \neg\alpha$ – is valid also for Priest's dialetheic view.[17] This is what he calls *classical* dialetheism, which opposes *intuitionistic* (or, better, *paracomplete*) dialetheism where valueless sentences are allowed. Though, there are some unusual, or maybe striking, consequences related to classical dialetheism. For example, «David Lewis's ninth wife did not have super powers» would be true (David Lewis had only one wife, Stephanie Lewis) and «This sentence is true» would be false. Despite being counter-intuitive, Priest accepts them as reasonably possible.

Finally, untruth. A sentence is untrue if it is not true, $\neg T\underline{\alpha}$. Let us examine the relation between truth and untruth first, and then that between untruth and falsity. Truth and untruth are exhaustive, since $T\underline{\alpha} \vee \neg T\underline{\alpha}$ is an instance of LEM. But they are not exclusive, because there are sentences which are both, e.g. an extended liar sentence such as «This sentence is untrue». Further, since truth and falsity are exhaustive, untruth entails falsity:

$$\neg T\underline{\alpha} \rightarrow T\underline{\neg\alpha} \tag{1.3}$$

However, the converse is rejected.[18] If $T\underline{\neg\alpha} \rightarrow \neg T\underline{\alpha}$ was accepted, falsity and untruth would collapse on each other – i.e., they would be the very same notion. They share two crucial features: both of them are exhaustive and nonexclusive with respect to truth. But Priest claims they are distinct: truth and untruth would be «"more inconsistent" than truth and falsity» (Priest, 2006b, p. 72). For truth and untruth would be both exclusive and nonexclusive, since $\exists x(Tx \wedge \neg Tx)$ and that from LEM can be also derived $\neg\exists x(Tx \wedge \neg Tx)$. On the contrary, truth and falsity would only be nonexclusive, since it seems impossible to argue that $\neg\exists x(Tx \wedge Fx)$. Of course, to say that truth and untruth are both exclusive and nonexclusive is a contradiction. This makes the dialetheic semantics inconsistent. But this is precisely what we are required to accept. As Priest emphasizes:

[17] A further reflection follows what we have just said. According to Priest, there are *negative* facts that make a sentence false: that there is no fact that makes that sentence true. But if we want to allow for sentences both true and false, we also need *positive* facts that speak in favor of the truth of the negation of a sentence. This is possible since we can have, for example, proofs of the negation of a sentence – as in the case of the liar paradox.

[18] But not, for example, in Beall's dialetheic view that validates $T\underline{\neg\alpha} \rightarrow \neg T\underline{\alpha}$.

> If I were attempting to produce a consistent theory of the inconsistent, this would be fatal. However, the aim of the enterprise is not to eliminate contradictions but to accommodate them.
>
> Priest (2006b, p.72)

1.2.2 Dialetheic semantics

Next step is to introduce the formal semantics that Priest considers the most suitable for dialetheism.[19] As we already revealed, a crucial point is to invalidate the principle of Explosion. That is, we need a paraconsistent (i.e., non-explosive) consequence relation. Since all this is well discussed in Priest (2006b, §§5-6 and §19), I will introduce such semantics fairly quickly, stressing a little more on the main passages.

Let us take an ordinary propositional language, $\mathcal{L}$, with the following (extensional) connectives: $\wedge$, $\vee$ and $\neg$. We can also define $\alpha \supset \beta$ as usual, i.e. $\neg\alpha \vee \beta$. We take P to be the set of propositional parameters and F the set of (well formed) formulas which corresponds to the closure of P under the connectives. Compared to classical formal semantics, the main change is in the evaluation function. In short, the functionality of the evaluation is dismissed. More precisely, we allow formulas to have up to two truth-values, which is not permitted if we take the evaluation to be a function. Nevertheless, from a technical point of view, the simplest way to do that is to preserve functionality and include a third truth-value, which corresponds to *both true and false*. Therefore, let ν an evaluation of sentence letter, i.e. a map from P to π, where $\pi = \{\{0\}, \{1\}, \{0, 1\}\}$. Now, we extend ν to range also over complex formulas by setting the following truth and falsity conditions:

($\neg_1$) $\quad 1 \in \nu(\neg\alpha)$ iff $0 \in \nu(\alpha)$

($\neg_2$) $\quad 0 \in \nu(\neg\alpha)$ iff $1 \in \nu(\alpha)$

($\wedge_1$) $\quad 1 \in \nu(\alpha \wedge \beta)$ iff $1 \in \nu(\alpha)$ and $1 \in \nu(\beta)$

($\wedge_2$) $\quad 0 \in \nu(\alpha \wedge \beta)$ iff $0 \in \nu(\alpha)$ or $0 \in \nu(\beta)$

($\vee_1$) $\quad 1 \in \nu(\alpha \vee \beta)$ iff $1 \in \nu(\alpha)$ or $1 \in \nu(\beta)$

($\vee_2$) $\quad 0 \in \nu(\alpha \vee \beta)$ iff $0 \in \nu(\alpha)$ and $0 \in \nu(\beta)$

As usual, we can read «$1 \in \nu(\alpha)$» as «α is true under ν» and «$0 \in \nu(\alpha)$» as «α is false under ν». Note that it is essential to specify also the falsity conditions, unlike in the

[19] The extensional propositional part of this logic is known as *the logic of paradox*, *LP* (or *LPQ* for its first-order quantified extension) and it was presented for the first time in Priest (1979).

case of classical logic. Logical truth and semantic consequence are defined in the standard way. Given Σ an improper subset of F,

$$\Sigma \models \alpha \text{ iff it is true of any evaluation, } \nu\text{, that if } 1 \in \nu(\beta) \text{ for all } \beta \in \Sigma \text{ then } 1 \in \nu(\alpha);$$

$$\models \alpha \text{ iff it is true of any evaluation, } \nu\text{, that } 1 \in \nu(\alpha).$$

It turns out that the classical two-valued logical truths are provably[20] all and only the logical truths of this semantics. Moreover, if α is a semantic consequence of a set of formulas, Σ, in this semantics, then it is also a classical two-valued semantic consequence of Σ, whereas the converse does not hold – e.g. the principles of Explosion and Disjunctive Syllogism (DS)[21] are not valid. (This will prove to be fundamental in discussing the classical recapture.)

Now, we need to extend such a dialetheic logic with quantifiers and identity. This can be done straightforwardly, in the usual way. There is just one novelty: the *anti-extension* for predicates. For every predicate P, its denotation is now represented by a pair of sets, $d_+(P)$ and $d_-(P)$, to be read as the extension of P and the anti-extension of P, respectively. $d_+(P)$ is the set of things in the domain (D) which satisfy P, whereas $d_-(P)$ is the set of things in the domain which satisfy the negation of P. In general, they are exhaustive (for all P, $d_+(P) \cup d_-(P) = D$) but not exclusive (there exists P such that $d_+(P) \cap d_-(P) \neq \emptyset$). All the remaining first-order logical machinery is the familiar one, except that we have to specify both truth and falsity conditions. For identity ($=$), the extension can be defined as $d_+(=) = \{\langle x, x\rangle \mid x \in D\}$, whereas the anti-extension is arbitrary, provided that $d_+(=) \cup d_-(=) = D$. As before, it turns out that all and only the dialetheic logical truths are those of classical first-order logic, and that if an inference is dialetheically valid then is also classically (first-ordered) valid – but the converse does not hold.

Two other notions need to be discussed before we can consider dialetheic semantics accomplished: the truth predicate and entailment. Let us begin with the first. We take the domain to contain all formulas and the vocabulary to include a set of constants, $\{\underline{\alpha}\}$, each of which denotes a unique formula α, for every formula of the language. Then, we designate the one-place predicate T as the truth predicate of the language. There are two conditions we require for T:

(T_1) $1 \in \nu(\alpha)$ iff $\alpha \in d_+(T)$

(T_2) if $\alpha \in d_-(T)$ then $0 \in \nu(\alpha)$

[20]See Priest (2006b, §5.5)

[21] $\{\alpha, \neg\alpha \vee \beta\} \models \beta$

where v is the usual evaluation function. (T_1) is essentially (1.1), the T-scheme, and (T_2) validates (1.3), which imposes exhaustivity to truth and falsity. We can also designate a one-place predicate, F, for the falsity predicate, and establish the two following conditions:

(F_1) $\quad 0 \in v(\alpha)$ iff $\alpha \in d_+(F)$

(F_2) $\quad$ if $\alpha \in d_-(F)$ then $1 \in v(\alpha)$

Finally, entailment. This is the notion at stake in «α entails β» and «if α, then it logically follows that β»; and it is fundamental since it is needed, for example, by semantics and set theory. It is widely assumed that the material implication is not appropriate to model entailment, but also that it is suitable for the ordinary conditional («if α then β» used as implication) and that its necessitation does the job of entailment correctly. However, this must be rejected in a dialetheic logic. For a minimal requirement for the conditional is that it satisfies the principle of *modus ponens* – i.e. $\{\alpha, \alpha \rightarrow \beta\} \models \beta$ – but in Priest's dialetheic setting the material implication does not do that, since DS does not hold. Therefore, the material implication is not the conditional, and for the same reason, its necessitation is not the entailment connective either. Because of that, he needs a different strategy. Besides, a further crucial remark concerns the Curry paradox: semantics based on a logic which contains the assertion principle – i.e. $(\alpha \wedge (\alpha \rightarrow \beta)) \rightarrow \beta$ – is trivial, i.e. everything is provable. Thus, an appropriate account of implication must not validate it.

Now, Priest changed his mind from the first edition of *In Contradiction* about what the best account of the conditional is. His most updated dialetheic conditional[22] is a relevant one and deploys impossible worlds in a Routley/Meyer semantics.[23] This is the account we are now going to introduce. As before, we will discuss the sentential relevant account first, and then add quantification and identity.

Let us take the sentential language of before and add an entailment operator, $\Rightarrow$. An interpretation is now a tuple $M = \langle P, I, R, @, v\rangle$. P is the set of possible worlds, and I is the set of impossible worlds – i.e. worlds where the laws of logic are different compared to those holding in $w \in P$ –, such that $P \cap I = \varnothing$ and $W = P \cup I$. @ is the actual world and is one of the possible worlds, $@ \in P$. R is a ternary accessibility relation on W.[24] v is an evaluation of the propositional parameters, i.e. a map $v : W \times Par \mapsto \pi$, where $\pi = \{\{1\}, \{0\}, \{1, 0\}\}$ and Par is the set of the propositional parameters. The way we extend v to complex formulas is the usual one. The truth and

[22] See Priest (2006b, p. 270-273).

[23] For an introduction to relevance logic, see Mares (2020).

[24] The philosophical interpretation of such a ternary relation is contentious – as it is, after all, for the notions of possible and impossible worlds. Various suggestions have been made about what R means, but none of them is entirely satisfactory. However, we can leave this matter aside here.

falsity conditions of the extensional connectives are the same as before, relativised to the appropriate world. For example, for every $w \in W$ (read $v_w(\alpha)$ as $v(w, \alpha)$):

($\wedge_1$) $1 \in v_w(\alpha \wedge \beta)$ iff $1 \in v_w(\alpha)$ and $1 \in v_w(\beta)$

($\wedge_2$) $0 \in v_w(\alpha \wedge \beta)$ iff $0 \in v_w(\alpha)$ or $0 \in v_w(\beta)$

Now, to distinguish between possible and impossible worlds we have to add the following constraint:

($*$) for all $w \in P$, $1 \in v_w(\alpha)$ or $0 \in v_w(\alpha)$

($*$) guarantees that LEM holds at every possible worlds; but since it is silent about I, LEM may fail at impossible worlds. Finally, the truth and falsity conditions for $\Rightarrow$:

($\Rightarrow_1$) $1 \in v_w(\alpha \Rightarrow \beta)$ iff for all $x, y \in W$ such that $Rwxy$, if $1 \in v_x(\alpha)$ then $1 \in v_y(\beta)$

($\Rightarrow_2$) $0 \in v_w(\alpha \Rightarrow \beta)$ iff for some $x, y \in W$ such that $Rwxy$, $1 \in v_x(\alpha)$ and $0 \in v_y(\beta)$

Given these conditions, it is easy to check that contraction[25] may fail: «[s]uppose that $w \in P$ and $p \Rightarrow (p \Rightarrow q)$ holds at w. Let x be any world such that p holds there; it follows that $p \Rightarrow q$ holds there. But it does not follow that q holds there: x may be a logically impossible world, so *modus ponens* may fail. Hence $p \Rightarrow q$ may not be true at w» (Priest, 2006b, p. 271). Also, this conditional verifies the usual implicational principles,[26] but it does not contrapose.[27] There is one problem, though: these conditionals do not satisfy LEM. We do not enter the details here, but suffice it to say that Priest fixes it by adding a further constraint he calls the *Augmentation Constraint*, and resorting to the presence of the trivial world – a world where every propositional parameter is both true and false – among impossible worlds.[28] Semantic consequence and logical truth are defined in the standard way:

> $\Sigma \models \alpha$ iff for all interpretations, M, it is true of the evaluation, v, that if $1 \in v_@(\beta)$ for all $\beta \in \Sigma$ then $1 \in v_@(\alpha)$.

[25] The principle of contraction is $p \to (p \to q) \models p \to q$, where $\to$ is an arbitrary undetermined conditional.

[26] See Priest (2006b, pp. 86-87).

[27] As Priest (2006b, p. 88) suggests, a contraposible conditional, $\alpha \to \beta$, can be easily defined as $(\alpha \Rightarrow \beta) \wedge (\neg\beta \Rightarrow \neg\alpha)$.

[28] See Priest (2006b, p. 272) and Priest (2006a, §5.2) for the detailed discussion.

$\models \alpha$ iff for all interpretations, M, it is true of the evaluation, ν, that $1 \in \nu_{@}(\alpha)$.

This semantics extends the truth-functional one so that everything we said before about logical truths and validity still applies.

Finally, we have to include quantification and identity. Priest opts for a constant domain quantified relevance logic. The (non-relevant) quantified language is augmented with the entailment operator, $\Rightarrow$. An interpretation is now a tuple $\langle P, I, R, @, D, d\rangle$, where P, I, R and @ are as before, D is the non-empty domain of quantification and d is the denotation function. s is any function which assigns a member of D (the same for every world) to each variable, and enable us to define the denotation of each term in the usual way. An evaluation, ν, is now a world-relativised function which maps a formula and s into π, satisfying the familiar recursive conditions. In the following, we show only the truth and falsity conditions for the entailment operator:

($\Rightarrow_1$) $1 \in \nu_w(\alpha \Rightarrow \beta, s)$ iff for all $x, y \in W$ such that $Rwxy$, if $1 \in \nu_x(\alpha, s)$ then $1 \in \nu_y(\beta, s)$

($\Rightarrow_2$) $0 \in \nu_w(\alpha \Rightarrow \beta, s)$ iff for some $x, y \in W$ such that $Rwxy$, $1 \in \nu_x(\alpha, s)$ and $0 \in \nu_y(\beta, s)$

In this first-order context, the condition $(*)$ becomes:

$(**)$ for all $w \in P$, $d_w^+(P) \cup d_w^-(P) = D^n$

where P is an n-place predicate, $d_w^+(P)$ and $d_w^-(P)$ are the extension and the anti-extension of P at w, and D is the domain of quantification. Regarding the identity predicate, it needs to satisfy the condition:

$(=_1)$ for all $w \in P$, $d_w^+(=) = \{\langle a, a\rangle \mid a \in D\}$,

whereas the anti-extension is arbitrary.

We now have a dialetheic full first-order logic with an entailment operator, and this formal apparatus is enough for our purposes.

1.2.3 Dialetheic pragmatics

To properly capture the overall dialetheic view, we still need to go into pragmatics, i.e. «the theory of the application of logic» (Priest, 2006b, p. 94). In particular, there is a couple of notions that play a very significant role in Priest's perspective: assertion and belief. We will follow Priest (2006b, §§4, 7) and Priest (2006a, §§4-6) to present his dialetheic pragmatics.

First, the speech act[29] of assertion. As we said in §1.2.1, we can not define assertion in terms of truth, on pain of circularity. But, according to Priest, for a good definition of assertion we just need to refer to Paul Grice's work.[30] Even if Grice did not explicitly attempt to define assertion, his notion of *non-natural meaning* can be easily applied to provide one:

> S asserts that p by the utterance u iff there is a hearer H such that:
>
> i. S intends u to produce in H the acceptance of p[31]
> ii. S intends H to recognize that i)
> iii. S intends H to accept that p at least partly for the reason that i)

Though, this account is not unproblematic. There are genuine examples of assertion that do not satisfy Grice's conditions. Nevertheless, Priest takes it to be a good first approximation which is suitable for his purposes. It is also important to define the dual notion of assertion: denial. This is because, as we will discuss in chapter 2, denial, together with Priest's accounts of rejection and negation, may have some undesirable consequences for dialetheism. Thus, we assume that:

> S denies that p by the utterance u iff there is a hearer H such that:
>
> i. S intends u to produce in H the rejection of p
> ii. S intends H to recognize that i)
> iii. S intends H to reject that p at least partly for the reason that i)

As it is clear, then, assertion and denial rely on the notions of acceptance and rejection. In this respect, Priest writes:

> Assertion and denial are [. . .] the linguistic expression of acceptance and rejection [. . .]. [T]he typical aim of assertion is to indicate that the utterer accepts the thing asserted, and, it may well be added, has appropriate grounds for doing so. (Derivatively, then, it often aims at getting the listener to accept it too.) [. . .] The typical aim of denial is to indicate that the utterer rejects the thing denied,

[29]For an introduction on speech acts see Green (2020)
[30]Grice (1957) and Grice (1968).
[31]A weaker version of this can be obtained by replacing «the acceptance of p» with «the acceptance that S accepts that p».

> and, again, one may add, has appropriate grounds for doing so. (Derivatively, then, it often aims at getting the listener to reject it too.)
>
> Priest (2006a, p. 104)

Many authors distinguish between acceptance and rational acceptance, and between rejection and rational rejection. From now on, we will be interested only in rational acceptance and rational rejection, and we will have them in mind even if we will be using just «acceptance» and «rejection» for the sake of simplicity.

Therefore, the propositional attitude of acceptance (or belief – throughout this work we will be using «to believe that *p*» and «to accept that *p*» as synonyms). As it is known, it is a very difficult task to set all the necessary and sufficient conditions for it. However, a sufficient condition for the acceptance of α, or a theory, is that we have good evidence (or good reasons) in support of it. Of course, this requires to define what a good evidence (good reason) is. This, too, is a challenging task. Nevertheless, suffice it to say that there are uncontroversial examples of good evidence in support of something: «that [it] can be deduced from something already rationally accepted; that it has experimental support; that it has high statistical probability, when this is all the information we have; and so on» (Priest, 2006b, p. 101). From this, Priest derives that it would be possible to believe a contradiction: we only need good reasons in support of it – which is essentially what he attempts to offer through his considerable work.
The dual notion of belief is rejection. A sufficient condition for the rejection of α, or a theory, is that there are good reasons against it. For example: «that it implies something we already have good reason to reject; that it is disconfirmed by the evidence; that it has a low statistical probability, where this is the only information we have; and so on» (Priest, 2006b, p. 102). Moreover, rejection and acceptance would not be exhaustive, since it is possible to be agnostic about α, or a theory. But Priest explicitly assumes them to be exclusive: «rational acceptance and rejection are mutually incompatible» (Priest, 2006b, p. 103). Thus, it would not be possible for a subject both to accept and reject α, whatever α is.

This matter is not so easy, though. In an attempt to further clarify such notions, Priest (2006a, p. 109-110) refines the sufficient conditions for acceptance and rejection as follows:

> **Accept**. One ought to accept something if there is good evidence for its truth.
>
> **Reject(U)**. One ought to reject something if there is good evidence for its untruth.

where the «ought» here is «one of rationality. It is rational to believe what is evidentially grounded (and irrational to believe what is not)» (Priest, 2006a, p. 110). A first remark is that such conditions are given in terms of (un)truth. This would seem to cause circularity. For Priest grounds truth in assertion, suggests defining assertion in terms of acceptance, and then makes use of (un)truth to express the norms for acceptance and rejection. A possible way out for him might be to emphasize that **Accept** and **Reject(U)** are not definitions and/or that it might be possible to rephrase such conditions without mentioning (un)truth. Be that as it may, these two conditions give rise to an important consequence. Consider, for example, «This sentence is not true». We have good evidence both for accepting and rejecting it. Therefore, rationality seems to force us to do both, even if this is arguably impossible. It may be suggested to replace **Reject(U)** with: one ought to reject something if there is good evidence for its untruth, unless there is also good evidence for its truth. However, this solution implies the failure of symmetry between truth and untruth (as well as between acceptance and rejection) and this is difficult to justify. Thus, Priest seems to hold both **Accept** and **Reject(U)**, but allowing for rational dilemmas:

> A dilemma is not a contradiction. Let us use the operator O, 'It is obligatory that', from standard deontic logic. Paradigm dilemmas are of the form: $O\alpha$ and $O\neg\alpha$, where α is a statement to the effect that something be done. More generally, in a dilemma there are two such statements α, β, such that $\Box\neg(\alpha \wedge \beta)$, yet $O\alpha$ and $O\beta$.
>
> Priest (2006a, p. 111)

Therefore, rational dilemmas are situations where rationality requires us to realise two incompatible propositional attitudes, i.e. to do the impossible. According to Priest, «the existence of dilemmas is simply a fact of life» (Priest, 2006a, p. 111). As it is accepted that there are moral and legal dilemmas, we should similarly accept that there are also rational dilemmas.[32] I will return to this point in chapter 2.

It is now convenient to examine how Priest takes denial and rejection to interact with negation. To see this, let us recall what is known as *the denial equivalence*:

> [. . .] denial and rejection should be understood in terms of negation, along with assertion and belief. [. . .] [T]o deny a content just is to assert its negation, and to reject a content just is to believe its negation.
>
> Ripley (2011, p. 622)

Frege (1952b) and Geach (1965) endorsed such a principle,[33] but Priest rejects it decisively. To see why, we need to start with negation. The classical negation – i.e. the negation connective as it behaves in classical logic – is exclusive. This means that the truth of α rules out the truth of $\neg\alpha$, and vice versa: a sentence and its negation

[32] Priest (2006a, §§6.6-7) argues that there are rational dilemmas regardless of dialetheism.

[33] But also Quine (2009) and Sorensen (2003), for example. For the denial equivalence is the mainstream view on the connection between acceptance/assertion, rejection/denial and negation.

are incompatible. On the contrary, the dialetheic negation is not exclusive: the truth of α and the truth of $\neg\alpha$ do not exclude each other, i.e. they are compatible. Because of that, it is possible for Priest to accept (and, consequently, to assert) both α and $\neg\alpha$. But now, if he accepted *the denial equivalence*, he would have to give up the exclusivity of acceptance and rejection (as well as that of assertion and denial). This move would be intolerable for him. For Priest needs at least one exclusive pair of notions to express his own view, i.e. to disagree with the non-dialetheic traditional view. Consider the Law of Non-Contradiction (LNC) in the following version: for every α, $\models \neg(\alpha \wedge \neg\alpha)$. Dialetheism claims that there are dialetheias, that is for some α, $\alpha \wedge \neg\alpha$. This is a counter-example which invalidates LNC. However, LNC holds in the dialetheic logic. If Priest assumed *the denial equivalence*, he would have both to accept and reject LNC, which gives no information about his position on LNC. Again, he wants to accept sentences like «The liar sentence is true» and «The liar sentence is false», and *not* (to be intended in an exclusive way) to reject them; similarly, he wants to reject «Trivialism is correct» and «Consistency is a necessary feature for a correct logic», and *not* (to be intended in an exclusive way) to accept them. To sum up: exclusivity of acceptance and rejection (and of assertion and denial) is necessary to express dialetheism, and arguably can not be avoided in a dialetheic account.

1.2.4 The classical recapture

Dialetheism extends classical logic[34] in the sense that it handles some inscrutable situations for classical logic, i.e. the inconsistent ones. However, it comes with a price: the invalidity of certain very intuitive rules of inference. Of course, intuition often proved to be misleading. But whenever we accept something counter-intuitive, we want to have an explanation to justify why our intuition leads us in a different direction. Take Special Relativity, for example. Time dilation and length contraction are thoroughly confirmed predictions of this theory, but they are highly counter-intuitive. However, there is a simple reason for that: for these *phenomena* to happen we need some uncommon physical conditions (i.e. very high speeds) that we do not directly experience in our everyday routine. In this regard, the most significant loss in the dialetheic account is DS: $\{\alpha \wedge (\neg\alpha \vee \beta)\} \models \beta$.[35] We appear to use it very often, both in our ordinary life and in the practice of highly reliable domains as those of mathematics and empirical sciences. Then, how should we make sense of its being invalid?

[34] It goes without saying that dialetheism flies in the face of what Field (2008) calls *Logical Dogmatism*: the idea that tinkering with classical logic is irrational since such a logic is indisputably superior to any other logic we can conceive. It's worth noting that logical dogmatists are not so many nowadays, and an increasing number of logician are getting into non-classical logics. To give just another example, Hartry Field considers paracomplete logic «to be our single all-purpose logic» (Field, 2008, p. 15) and the correct answer to 'save truth from paradox'.

[35] Note that the dialetheic-invalidity of other very important classical inferences, such as *modus ponens*, *modus tollens* and the *reductio ad absurdum*, strictly depends on the lack of DS.

In the first place, according to Priest (2006b, p. 111) it must be acknowledged that a counter-model to DS requires α to be a dialetheia, that is an inconsistent situation: to get $\alpha \wedge (\neg\alpha \vee \beta)$ true and β false, we need $\alpha \wedge \neg\alpha$ to be true. Therefore, we may be tempted to implement the following strategy:

1. express consistency through a sentence of the formal language $\mathcal{L}$;
2. impose it as an additional premise of DS to force the situation to be consistent, and get back its validity.

Unfortunately, this strategy fails because of the impossibility of step 1: there is no way we can express consistency in $\mathcal{L}$. Thus, Priest's first approach is to impose a DS-like principle at the pragmatic level. This is what he names *principle R*: «If a disjunction is rationally acceptable and one of the disjuncts is rationally rejectable, then the other is rationally acceptable» (Priest, 2006b, p. 113). However, if there are rational dilemmas – as Priest (2002, §6.5) claims – this principle cannot hold. For let α be a rational dilemma, that is both acceptable and rejectable (in theory, but not in practice). For any β, α entails $\alpha \vee \beta$, so this is acceptable too. By principle R, β would be acceptable. Therefore, everything would be acceptable, which obviously must be rejected. The dialetheist is then required to implement a different strategy.

Following Priest (2006b, p. 110), let say that «an inference [is] *quasi-valid* if it involves essentially only extensional connectives and quantifiers, and is classically valid but dialetheically invalid» and call the contradiction that, if true, invalidates DS, $\alpha \wedge \neg\alpha$, the *crucial contradiction*. The principle R was precisely intended to justify the use of quasi-valid inferences in consistent situations. Now, according to Priest, there would be another reason that allows the recapture of classical logic: the low *a priori* statistical frequency of true contradictions. With Priest's words: «[t]he normal success of quasi-valid reasoning [...] provides the basis of a transcendental argument for the infrequency of dialetheias» (Priest, 2006b, p. 116). Thus, consistency can be taken as a default assumption which justifies the use of classically valid inferences in consistent situations. That is, it seems we are justified in assuming consistency until and unless it is shown otherwise. All this leads to the endorsement of the following *Methodological Maxim*: «[u]nless we have specific grounds for believing that the crucial contradictions in a piece of quasi-valid reasoning are dialetheias, we may accept the reasoning» (Priest, 2006b, p. 116).

Priest (2006b, §16) makes use of this idea to develop a formal theory of reasoning that he calls *Minimally Inconsistent LP*, or *LPm*, which encapsulates the classical recapture. His strategy is well summarized in the following passage:

> [...] given some information from which we have to reason, we can cash out the idea that the situation is no more inconsistent than we are forced to assume

> by restricting ourselves to those models of the information that are, in some sense, as consistent as possible, given the information – or, as we will say, are minimally inconsistent.
>
> Priest (2006b, p. 222)

We do not go into technical details here. Suffice it to say that this formalisation is possible and requires the following steps: (1) to define an appropriate measure for the consistency of a model, so that a consistency ordering of models can be obtained, (2) to define the notion of *minimally inconsistent model* (m.i. model) and (3) to define a new consequence relation based on the notion of m.i. model. In short, step (1) is achieved by defining the consistency-degree of a model $\mathcal{M}$ as the set $\mathcal{M}!$ of atomic facts with value $\{1, 0\}$ in $\mathcal{M}$. Therefore, given any two models $\mathcal{M}_1$ and $\mathcal{M}_2$, $\mathcal{M}_1$ is more consistent than $\mathcal{M}_2$ iff $\mathcal{M}_1!$ is a proper subset of $\mathcal{M}_2!$. For step (2), we can define the m.i. model of Σ[36] as the interpretation $\mathcal{M}_{mi}$ such that, given any other interpretation $\mathcal{M}'$ of Σ, if $\mathcal{M}'$ is more consistent than $\mathcal{M}_{mi}$ then $\mathcal{M}'$ is not a model for Σ. Finally, we can say that a formula α is a minimally consistent consequence of Σ iff every m.i. model of Σ is also a model of α (step (3)). Thus:

> *LPm* is a more generous notion of consequence than *LP*, which allows for classical inferences — such as the disjunctive syllogism — provided inconsistency does not "get in the way"; in particular, it is identical with classical logic in consistent situations. It thus gives a precise account of how it is that classical inferences are acceptable, paraconsistently, in consistent situations.
>
> Priest (2006b, p. 225)

To conclude, dialetheism together with the low probability of dialetheias and the Methodological Maxim recaptures the whole power of classical logic, and extends it to handle inconsistent situations.[37]

1.3 Main advantages

According to Priest, we should accept dialetheism because it is more cost-effective than its competitors. To be more precise, he claims that we should go for dialetheism because it is the most rational choice we can make to properly account for some relevant *phenomena* – e.g. logical paradoxes of self-reference, transition states, inconsistent laws, etc. This means that it would be rationally preferable to all the other non-dialetheic (consistent) theories. Then, recalling what we said in §1.1, we would

[36] Σ is an arbitrary set of formulas.

[37] Such result can be interpreted as a *prima facie* good score of dialetheism with respect to the *criterion* of how it coheres with the extant net of knowledge.

have:

$$K_d \sqsubset K, \quad \text{for every } K \text{ such that } K \neq K_d$$

where K_d represents dialetheism and K runs over every consistent (or trivial) theory. This is to say that the weighted average of the $\mu_{c_i}(K_d)$, $\rho(K_d) = [\rho^-(K_d), \rho^+(K_d)]$, is greater than $\rho(K)$ for every K such that $K \neq K_d$. Thus, we would need that $\rho^-(K_d) > \rho^+(K)$ for every K other than K_d. However, this is far from being accepted by the majority of those who work in this field. It is not at all clear whether dialetheism is the most rational choice, and to be fair it is still rejected by the majority of philosophers. In this section, I will briefly present its main advantages, whereas in the next I will display the criticisms that the literature has brought to light.

1.3.1 Logical paradoxes of self-reference

Arguably, the logical paradoxes of self-reference are the main reason why dialetheism might be very appealing. A paradox is «an apparently unacceptable conclusion derived by apparently acceptable reasoning from apparently acceptable premises» (Sainsbury, 2009, p. 1). More precisely, it is a seemingly sound argument ending in a contradiction. The paradoxes of self-reference are those which stem from the mechanism of self-reference, where a sentence contains its own name or a group of sentences generate a closed circuit of reference.[38] Usually, logical paradoxes are considered problematic so much so that one of the tasks of the logician is to dismantle them by showing what is wrong with the argument, primarily through the identification of the deficiency in our understanding of the central concepts (e.g. truth, set, etc.) involved in it. However, although (some of) the paradoxes of self-reference have been known for more than nearly two-and-a-half millennia, there is no agreed-upon solution for them.[39]

This family of paradoxes is generally divided into two categories: semantic paradoxes and set-theoretic ones.[40] Semantic paradoxes involve linguistic concepts such as truth, denotation, definability, etc. Some well-known examples are the liar paradox and its strengthened forms, Grelling's, Berry's, Richard's and Köenig's.

[38]For an introduction on self-reference see Bolander (2017).

[39]For an overview of recent developments in approaches to solving these paradoxes see Murzi and Carrara (2015).

[40]This partition dates back to Ramsey (1926). Nowadays, it is commonly accepted that paradoxes from both these families are strictly related, if not the consequence of the very same *phenomenon* (see e.g. Priest, 1994). Therefore, they might be conceived as just one class of paradoxes (e.g. the *inclosure paradoxes*). However, this partition is still in use because of its convenience. Note that there is also another family of paradoxes very close to logical paradoxes: the epistemic paradoxes of self-reference. These involve epistemic concepts such as knowledge and belief and could be treated alike logical paradoxes. For an introduction see Sorensen (2020).

Tarski (1936) identifies the cause of semantic paradoxes in the semantic closure[41] and from this analysis some alleged solutions have been proposed. These follow one of three main strategies. The first strategy is to deny that self-referential sentences are meaningful (see e.g. Pleitz, 2018). The second strategy is to allow for neither true nor false sentences by dismissing LEM, i.e. *gap* truth-values (see e.g. Kripke, 1976 and Field, 2008). The third strategy is to limit the expressive power of the natural language which is used to express the paradox: English, for example, would be a hierarchy of semantically open languages such that each of them has a truth predicate which can be legitimately applied only to the sentences of the languages below. Unfortunately, none of these strategies appear to be entirely convincing. For self-referential sentences appear to be meaningful, the "gappy" solutions suffer from the revenge of many forms of strengthen liar paradox,[42] and a hierarchical theory of truth cannot be true according to its own account[43] – besides the fact that natural languages do not seem to show such a hierarchical structure.[44]

Set-theoretic paradoxes involve set-theoretic concepts such as membership, cardinality, etc. Some examples are the Russell's paradox, Cantor's, Burali-Forti's and Mirimanoff's. They are strictly related to the naive notion of set and in particular to the principle of abstraction[45] which leads to inconsistency. Therefore, the usual solution – though not the only one, e.g. Russell's *type theory* – is to reject this principle in favor of the axiom of separation, a restricted version of abstraction which avoids contradictions. However, this move produces the so called *cumulative hierarchy* of sets: only the sets belonging to such a theoretic structure would exist. But the claim that the cumulative hierarchy exhausts the universe of sets is not so obvious[46] and some (e.g. Priest) take it to be a too high cost to elude inconsistency.

That said, dialetheism offers a simple but unorthodox solution to logical paradoxes. According to Priest, they are what they *prima facie* seem to be: sound arguments. Therefore, their contradictory conclusions are true, i.e. they are dialetheias. We do not need to impose any infinite *ad hoc* stratification in our natural languages,

[41]This is a set of closure conditions such that if a theory satisfies all of them is inconsistent. For a simple discussion see Priest (2006b, p. 11).

[42]This is not the only problem with gappy solutions. A further crucial problem is that a gap theory cannot affirm that gaps are gap. About that, see Priest (2006b, §3) and Beall, Glanzberg, and Ripley (2018).

[43]In this regard, I quote Weber (2021, p. 10)'s words : «The most important problem for any hierarchy, though, is the question: how can a hierarchical theory be *true*, according to itself? True claims must be indexed to some level of the never-ending hierarchy, but the claim "all true claims must be indexed to some level of the hierarchy" cannot be so indexed» (emphasis in original).

[44]About the lack of such a stratification in natural languages, consider the compelling examples given in Kripke (1976, p. 692).

[45]Informally, given any condition (i.e. well-formed formula) β, the principle of abstraction guarantees the existence of a set including all and only the objects satisfying β.

[46]Actually, the majority of mathematicians consider this universe of sets the correct and natural one. For example, Myhill (1984) reports that Kurt Gödel once remarked: «There never were any set-theoretic paradoxes». Field (2008, p. 3) gives the following interpretation of Gödel's observation: «The idea behind [...] his remark is presumably that the notion of set was hierarchical from the start, so that it should have been obvious all along that there was no Russell set».

and we can still rely on LEM. We have just to accept that sentences like «This sentence is false» are both true and false, i.e. to accept inconsistency (but not triviality). And for what concern set-theoretic paradoxes, they would just prove the existence of inconsistent but legitimate sets which do not fit the cumulative hierarchy (e.g. the Russell's set). As a matter of fact, some interesting non-trivial inconsistent set theories have been developed.[47] Moreover, if we accept that the semantic and the set-theoretic paradoxes both share the same structure,[48] dialetheism satisfies what Priest (1994, §5) calls the Principle of Uniform Solution (PUS): «same kind of paradox, same kind of solution». Thus, PUS would represent a further reason in support of the dialetheic solution.

To conclude, dialetheism would be rationally preferable to the other solutions with respect to the *criteria* of simplicity – since it solves both semantic and set-theoretic paradoxes all at once – and explanatory power – since it makes the problems with logical paradoxes disappear.[49]

1.3.2 Transition states

Being related to semantics and set theory, the dialetheias produced by logical paradoxes are quite abstract, in the sense that they concern domains of discourse including abstract objects such as sets, sentences, etc. However, according to Priest there would be other contradictions affecting the empirical world. This would be the case, for example, of transition states.

A transition state is the condition of an object at the instant of time in which some of its properties change. More precisely, it is the state of a system s at instant t_0 such that, before t_0 s is in a state correctly described by α, and after t_0 s is in a state correctly described by $\neg\alpha$. The most representative case is that of motion, where the property that changes is the location of the moving object with time. For example, at the time t_1 the moving object is located at x_1, whereas the same object is not located at x_1 at the time t_2, since it has moved. Now, consider the following example given by Priest:

---

[47] On this topic see for example Routley (1979), Brady (1989) and Weber (2012).

[48] This is what Priest (1994) contends. The common structure would be the so called *Inclosure Schema*. Given two predicates P and Q, and a possibly partial function δ, the Inclosure Schema consists of the following two conditions:

1. $\omega = \{x : P(x)\}$ exists and $Q(w)$ holds;
2. if y is a subset of w such that $Q(y)$ holds then:
   (a) $\delta(y) \notin y$;
   (b) $\delta(y) \in y$

If these conditions are satisfied we have that $\delta(w) \notin w$ and $\delta(w) \in w$ (i.e. a contradiction) by 2a and 2b, since w is trivially a subset of w and since $Q(w)$ holds by condition 1.

[49] For a general critique of such a view see Field (2008, §23-26). For a dissenting analysis of the preference of the dialetheic solution based on PUS see Smith (2000).

> I am in a room. As I walk through the door, am I in the room or out of (not in) it? To emphasize that this is not a problem of vagueness, suppose we identify my position with that of my centre of gravity, and the door with the vertical plane passing through its centre of gravity. As I leave the room there must be an instant at which the point lies on the plane. At that instant am I in or out?
>
> Priest (2006b, p.161)

More generally, what is the correct description of the system at t_0, i.e. in the transition state? α? $\neg\alpha$? Their conjunction? Neither? Classical logic requires that the answer must be either α or (to be read in an exclusive way) $\neg\alpha$. However, because of the continuity of time this conclusion is highly disputable. For if time is continuous, any property that holds at any continuous set of times holds at any temporal limit of those times. Therefore, if we consider two consecutive continuous set of times, in the transition point (the point shared by the two sets) both of the properties hold. Here is where dialetheism comes into play.

The dialetheic solution is to accept that the correct description of the transition state is $\alpha \wedge \neg\alpha$, i.e. a dialetheia. For example, we have to accept that «something is a cup and not a cup the instant it breaks into pieces» (Priest, 2006b, p. 170). To this regard, a relevant case is that of the arrow paradox, one of Zeno's paradoxes. These paradoxes are as old as the Liar, but their history is quite different. The paradoxes were much discussed in Ancient Greek philosophy and Medieval philosophy. However, it is usually claimed they were finally solved by developments in 19th Century mathematics. For the most part, this is right. However, this is arguably not the case for the arrow paradox. We can formulate this paradox as follows. Take an arrow – or a point particle, to avoid irrelevant worries – in motion from a to b. At any instant of its motion, the progress made by the particle in its journey is zero, since this is just an instant. But the time of flight is composed of such instants. So the progress made on the journey is the sum of the progresses made at each instant. But zero plus zero plus ... as many times as you like – even infinitely many times – is zero. So the particle makes no progress on its journey at all: it does not move. The standard solution to this paradox is just to bite the bullet. The particle makes no progress at each instant, but somehow, in the sum of instants, it does.

According to Priest, dialetheism provides a more illuminating answer. Suppose that at some instant, t, the particle is at position x. Then, since it is in motion, it is also at a place a little bit after it, say $x + \epsilon$; and also a place a little bit before it, say $x - \epsilon$. Since it is at all these places, it does make progress at t. Hence, it can make progress in the sum of all instants. This solution implies that the particle is in a contradictory state at every instant of time during its motion. For since it is at $x + \epsilon$ it is not at x, even though it is. In fact, for the particle to be in motion would be for it to realise such a contradictory state. If it were not in motion at t, it would simply be at x. Strange as it may be, this solution is very close to the way Hegel conceived the motion and, according to Priest, it would correctly address the problems with transition states.

1.3.3 Inconsistent law

According to Priest, the most transparent examples of dialetheias would be those which surface from inconsistent legal systems. If we accept their existence, consistent theories cannot make sense of them. In this regard, dialetheism would be rationally preferable with respect to its consistent competitors because of its greater explanatory power.

Laws are the kind of statements that can be made true simply *by fiat*.[50] Many things cannot be made true *by fiat* (e.g., that the Moon is more than a kilometer from the Earth, that the Sun is shining, etc.); but duly constituted legislature can make some things the case, simply by passing the appropriate legislation (e.g., that people in a certain class have a legal right or duty). Now, consider a situation in which some University Library Regulations include the two following rules:

R_1) Any professor of the University can access the library any day of the week, (only) from 10 a.m. to 10 p.m.

R_2) Any student of the University can access the library any day of the week, (only) from 10 a.m. to 8 p.m.

Therefore, from 8 p.m. to 10 p.m. professors are allowed to enter the library, but the students are not. We may suppose that at the time when the University Library Regulations were passed, the possibility that there might be a professor who was also a student was unthinkable. But now, suppose that the renowned archaeologist and university professor Hershel Layton becomes passionate about mathematics, and enrolls in his own University as a maths student. Therefore, he is both a professor and a student, and according to the University Library Regulations he can and can not enter the library on Monday at 9 p.m. That, of course, is a contradictory situation. Now, in many jurisdictions there are some standard procedures for resolving contradictions of this kind. For example, laws may be ranked in increasing order of weight: precedence law, statute law, constitutional law. And where a law from one level contrasts with one of higher order, it is the one of higher order which prevails. Or again, in some jurisdictions, there is a principle of *lex posterior*, according to which, if an early law conflicts with a later law, the later one takes precedence. But we may suppose that none of these mechanisms applies in the present case: both rules were in the same piece of regulations, passed at the same time, and so on. Then the contradiction occurs. Of course, if this situation were to arise, the regulations

[50] It is contentious whether or not a legal statement bears a truth-value. Priest (2006b, §13.2) argues it does.

would be changed. The function of the law (or rules) is a very practical one, and the contradictory situation is not practical. But this does not alter the fact that before the change the situation is contradictory. That, indeed, is why the change is necessary.[51]

1.3.4 Dialetheism as a fruitful ground to develop metaphysical theories

Another important advantage of dialetheism is that it would represent the base upon which to develop further theories capable of solving long-standing metaphysical problems. For instance, this would be the case of *gluon theory* that was developed by Priest (2014b) to solve *the problem of the one and the many*.[52] Of course, the evaluation of such a fruitfulness of dialetheism depends on the strength and correctness of the metaphysical theories it allows to conceive. More precisely: the assessment of dialetheism with respect to this point depends on being the other theories based on it rationally preferable than their competitors. Thus, the evaluation of such theories is necessary.

1.3.5 Other suggested advantages

There are other alleged advantages put forward by some dialetheists. But since they are (even) more contentious and/or less discussed, I will limit myself to list them summarily.

Vagueness and the sorites paradox

Vagueness[53] is a highly investigated topic, and the same is for a paradox that is generated from vague terms: the sorites paradox.[54] Some strategies to deal with vagueness have been proposed, such as many-valued logics, supervalutationism and contextualism. Unfortunately, there is no general agreement about what the correct account of this *phenomenon* is, and the same goes for the solution of the sorites paradox. An alternative approach is subvalutationism, the dual of supervalutationism. The key idea is to analyze borderline cases in terms of truth-value gluts. Therefore, a dialetheic semantic is required and this would be further evidence for dialetheism.[55]

The limits of thought

According to Priest (2002), there would be another sort of situation which provides an important application of dialetheism: one concerning the limits of thought. There are many theories in the history of philosophy according to which there are things

[51] For a recent discussion, see Beall (2016).
[52] I postpone its discussion to §3.
[53] For an introduction on vagueness see Sorensen (2018).
[54] For an introduction on the sorites paradox see Hyde and Raffman (2018).
[55] On this topic see e.g. Hyde (1997), Priest (2010) and Weber (2010).

(e.g. God, the ultimate reality, the form of propositions, etc.) which are beyond our ability to describe or conceptualise. Of course, even to say that there are things of this kind is to describe/conceptualise them. So there appears to be a contradiction here. Instead of trying to avoid the contradiction, Priest's suggestion is to embrace it. This means to accept that these ineffable objects are dialetheic, i.e. their correct descriptions are inconsistent.

Contradictory psychological states

Some psychological states appear to be contradictory, for example the state experienced by someone who finds something so repulsive that is compelled to look at it. She would be attracted and repelled at the same time. Therefore, *prima facie* that is a contradiction. Priest (2019) argues it is a true one. Thus, dialetheism would make sense of such contradictory psychological states.

1.4 Main criticisms

Let us now turn to arguments against dialetheism. One may object to a possible application of dialetheism, such as those in the previous section, on the ground that there are better and consistent ways of handling the matter at issue. Let us call these *local* objections. But one may object not simply to a particular application of dialetheism, but to dialetheism as such. Call these *global* objections.

Now, what would be an effective global objection to dialetheism is a convincing defence of LNC. Unfortunately, nobody has never been able to give a reliable argument in support of it. The major attempt was that of Aristotle, (see *Metaphysics*, book Γ, 3-6), but his justification is highly problematic and unsatisfactory.[56] David Lewis suggested that it would not even be possible to give such a justification:

> [t]o conduct a debate, one needs common ground; principles in dispute cannot of course fairly be used as common ground; and in this case, the principles not in dispute are so very much less certain than non-contradiction itself that it matters little whether or not a successful defence of non-contradiction could be based on them.
>
> Lewis (2004)

> Nothing is, and nothing could be, literally both true and false. This we know for certain, and *a priori*, and without any exception for especially perplexing subject matters. [. . .] That may seem dogmatic. And it is: I am affirming the very thesis that [the rivals of the Law of Non-Contradiction] have called into question and – contrary to the rules of debate – I decline to defend it. Further, I

[56] For a critique of Aristotle's defence of LNC see Lukasiewicz (1971), Dancy (2012) and Priest (2006a)

> concede that it is indefensible against their challenge. They have called so much into question that I have no foothold on undisputed ground. So much the worse for the demand that philosophers always must be ready to defend their theses under the rules of debate.
>
> Lewis (1982, p. 434-5)

Be that as it may, there are other modern criticisms that have been put forward against dialetheism. In what follow I will present the main ones.

1.4.1 Objections based on the lack of exclusive notions

This kind of global criticism is arguably the most frequent, if not the most pressing and problematic for dialetheism. It involves the concept of *exclusion* (or, equivalently, that of *incompatibility*) in so far as dialetheic semantics lacks any exclusive notion. It is no coincidence, in fact, that this argument is known as the *exclusion problem*.[57]

It comes in some slightly different versions,[58] but all of them boil down to the very same idea. Roughly, a notion is said to be exclusive when it rules out something. As we have already discussed in §1.2.3, classical negation is exclusive (the truth of α excludes the truth of $\neg\alpha$), whereas dialetheic negation is not (the truth of α is compatible with the truth of $\neg\alpha$). Furthermore, in the dialetheic semantics there is no other exclusive notion available, and this appears to be problematic. For an exclusive notion seems to be necessary for at least the following tasks: to disagree with others, to formulate meaningful sentences and convey information, and also to express the very same view of dialetheism.

For what concern disagreement, the problem is well presented by Parsons (1990) in this famous passage:

> Suppose that you say "β" and Priest replies "$\neg\beta$". Under ordinary circumstances you would think that he had disagreed with you. But then you remember that Priest is a dialetheist, and it occurs to you that he might very well agree with you after all – since he might think that "β" and "$\neg\beta$" are *both* true. How can he indicate that he genuinely disagrees with you? The natural choice is for him to say "β is not true". However, the truth of this assertion is also consistent with β's being true - for a dialetheist, anyway. So [...] Priest has difficulty asserting disagreement with other's views.
>
> Parsons (1990, p. 345)

[57]Sometimes it is also referred as the *just true problem*, e.g. Rossberg (2013). However, for the sake of completeness, note that Young (2015) has recently argued that these are two different problems.

[58]In this respect, some relevant works are Parsons (1990), Littmann and Simmons (2004), Shapiro (2004) and Berto (2006).

The same point has been also emphasized by others, e.g. Shapiro (2004) and Littmann and Simmons (2004). For example:

> One can refute other opponents by showing that their views lead to contradiction, especially if the inconsistency comes by the opponent's own lights. The debate usually turns on whether the view does in fact entail a contradiction. This cuts no ice against the dialetheist, since she embraces contradictions. For her, a *reductio ad contradictionem* is not a *reductio ad absurdum*.
>
> Shapiro (2004, p. 337)

The moral is that of Popper (1940): a critical dialogue between a dialetheist and a non-dialetheist (where «non» is meant as exclusive) seems to be hopeless.

Another way to put the exclusion problem is to say that if we accept dialetheism we cannot explain how communication is possible. For a sentence is meaningful if it is capable of expressing a content; that is, if it can draw a distinction between the situations (facts, state of affairs, worlds, etc.) where it holds and those where it does not. But a dialetheia cannot do that. And since in the dialetheic theory every sentence is potentially both true and false (due to the theory-metatheory indistinguishability), no information could be properly conveyed. After all, «[o]nce we realize that the theory includes not only the statement "(L) is both true and false" but also the statement "(L) isn't both true and false" we may feel at a loss» (Littmann and Simmons, 2004, p. 318). A dialetheist may reply that dialetheism does convey information on the grounds that the supporter of the LNC argues against it – so she must have understood it. However, what she may understand is just a consistent subset of the sentences of the theory, and not what the theory as a whole tells about the world.

Finally, because of the lack of exclusive notions dialetheism is said to be unable to express its own view. Recall that, according to Priest, there is no object-language/metalanguage semantic distinction, and dialetheic semantics is all you need (see §1.2.1). That is, when we talk about dialetheism, we use the very same dialetheic semantics – including the non-exclusive negation – to discuss it. So, let us consider some crucial dialetheic claims:

> Notice also that the rational acceptability and rejectability of something, though *not* exhaustive, are certainly *in*compatible.
>
> Priest (2006b, p. 103, my italics)

> Choosing is an irredeemably goal-directed activity. And as we have seen, such action is *in*compatible with believing everything. It follows that I can*not* but reject trivialism. Phenomenologically, it is *not* an option for me.
>
> Priest (2006a, p. 70, my italics)

> The *Non*-Triviality of the World
>
> Title of §6.4 of Priest (2006a, my italics))

> English certainly does *not* seem to be of this form. Its "surface" structure is certainly *not* of this form [...]. There is [...] *no* linguistic or grammatical evidence at all that the English predicate "is true" does typically ambiguous duty for an infinite hierarchy of predicates at the deep level.
>
> Priest (2006b, p. 19, my italics)

> [...] there are *no* truth value gaps.
>
> Priest (2006b, p. 13, my italics)

Now, since Priest employs a non-exclusive negation to express such claims, they are compatible with their contraries. But, for example, since he rejects trivialism he want to say that *only* some contradictions are true, or equivalently that there are contradictions that are *only* false. Therefore, he does seems to require an exclusive notion to describe his own position.[59]

As far as I know, five main strategies have been proposed so far to cope with the exclusion problem. The first one is Priest's approach. He tries to stand his ground by recovering exclusivity at the pragmatic level. He resorts to the pairs of notions we have already discussed in §1.2.3: acceptance/assertion and rejection/denial. Since Priest takes them to be exclusive, he argues that a dialetheist can actually do what it is said to be unable to do by means of them. For example, he can express disagreement by rejecting and denying his opponent's position – e.g. trivialism. Or he can discuss and defend dialetheism by accepting and asserting the claims we have previously reported. In other words, the sentence «There are no truth value gaps» is not enough to express the exclusion of gaps, but its assertion would do the job. However, this response does not appear to be entirely convincing. For pragmatics operators suffer from well-known expressive limitations.[60] Further, in chapter 2 I am going to present a new damaging consequence of Priest's pragmatic account – in the form of a revenge paradox – which undermines its reliability.

The second strategy is *arrow-falsum*, $\rightarrow \bot$. This is an exclusion-expressing device which can be defined using the dialetheic semantics: $\bot$ (*falsum*) is a logical constant such that it is a logical truth that $\bot \rightarrow \alpha$, for every α (i.e. $\bot$ represents something unacceptable also for a dialetheist), and $\rightarrow$ is a detachable conditional. Thus, the dialetheist may try to rule out α by uttering $\alpha \rightarrow \bot$. But *arrow-falsum*

[59]It is worth noting that, as Berto (2012) points out, the exclusion problem has a formal counterpart: «the semantics of various paraconsistent logics, such as the standard dialetheist's favourite one, LP, admits a so-called trivial model, or trivial interpretation. [...] It seems, thus, that nothing is ruled out on logical grounds only in the standard dialetheic framework. If something is just plainly untrue, this has to be settled on non-logical grounds» (Berto, 2012, p. 176).

[60]See §2.1, as well as Shapiro (2004, pp. 339–340) and Field (2008, pp. 387–388).

rules out sentences based on the implication of something unacceptable, and does not capture the more simple way of ruling out something grounded on empirical facts. Moreover, Field (2008, pp. 388-389) has showed that *arrow-falsum* can only be a partial exclusion-expressing device. For instance, take the standard Curry sentence:

(*k*) $T\underline{k} \rightarrow \bot$

If (*k*) was true we would get $\bot$. Therefore, we want to rule it out. But we cannot do this by uttering $k \rightarrow \bot$, for this just is (*k*) and we would get $\bot$ again.

The third approach consists in restoring the theory/metatheory distinction and going for a consistent metatheory with an exclusive negation. This view has been proposed by Rescher and Brandom (1980), but it goes without saying that Priest firmly rejects it for the following two reasons: the reappearance of revenge paradoxes in the metatheory and the lack of evidence of such a distinction in natural languages. With his words:

> It has been felt by some that, even if our object-theory is inconsistent, our metatheory should be consistent [...]. [...] I reject this view categorically. [...] [T]his distinction is a spurious one based on incorrect attempts to impose consistency. A natural language (or a formal language that models that aspect of its behaviour) can give its own semantics [...] [that] is a paradigm example of an inconsistent area.
>
> Priest (2006b, p. 70)

The fourth option is to rely on Gricean implicature.[61] Given the conversational maxim to say all that is relevant, if I say α and a dialetheist, e.g. Priest, replies with $\neg\alpha$, I might conclude that he disagrees with me. For if he thought that $\alpha \wedge \neg\alpha$ were true, he would have said so. However, things are not so simple. As Shapiro (2004) points out, we could have the following situation:

> [...] I assert β and Priest replies with $\neg\beta$ [...]. Suppose, however, that he believes $\neg\beta$ and is unsure whether β is also true. Then he would not assert $\beta \wedge \neg\beta$, since he does not believe that, but he does not quite disagree with me either. So for a dialetheist, the bare assertion of $\neg\beta$ does not carry the implicature that he disagrees with β.
>
> Shapiro (2004, p. 339)

[61] For an introduction on the topic of implicatures see Davis (2019).

In this case, to escape the problem we might take the conversational maxim to require to express Priest's non-disagreement, e.g. by saying «$\neg\beta$, but I am not sure about β». However, it is far from clear whether or not this approach is truly satisfactory.[62]

Finally, the last strategy is to take the notion of *exclusion* as primitive and use it to improve the expressive power of dialetheic semantics. This solution has been proposed by Berto (2014). He introduces the symbol $\bullet$ for the primitive exclusion relation which can be taken to hold between properties. Thus, $P \bullet Q$ is to be read as «Properties P and Q are incompatible», or as «Being P rules out being Q». Then, he defines a predicative underling functor, _ (not to be confused with the same symbol we have been using so far as a quotation device), such that «[t]aking a property P as input, [it] outputs its minimal incompatible $\underline{P}$, the having of which is the having of a feature ruling P out» (Berto, 2014, p. 7):

$$\text{(Def\_)} \quad \underline{P}x \stackrel{\text{def}}{=\!=} \exists Q\,(Qx \wedge P \bullet Q)$$

This allows us to define what Berto calls *absolute contradiction*: $Px \wedge \underline{P}x$. Unlike $Px \wedge \neg Px$, also for the dialetheist the absolute contradiction should hold for no P and no x. Therefore, it seems we have got back Lewis' requested common ground for debating dialetheism: «a notion of contradiction [...] unacceptable by any involved party» (Berto, 2014, p. 10). However, the matter is not so simple and Berto is well aware of this. For with his words:

> Priest want the metatheory to be itself dialetheic [and h]ow this relates to the exclusion-expressing problem depends on how the dialetheic set-theoretic framework is developed. [...] A verdict on this issue should wait for precise applications of the [...] inconsistent theories of sets to dialetheic theories of truth.
>
> Berto (2014, p. 9, fn. 14)

1.4.2 Curry's paradox

Curry's Paradox represents a contentious point about the dialetheic approach. There are essentially two reasons for that. The first one is that stronger forms of Curry's paradox are not blocked by the strategy we have discussed in §1.2.2.[63] This may lead to contraction-free substructural logics where principles other than contraction (e.g. transitivity) are dropped.[64] The second is that some believe that also the Curry's paradox share the same structure of the other logical paradoxes.[65] Consequently, in

[62]Consider, for example, Shapiro (2004, p. 340). Here, the author is sceptical about how implicature can help the dialetheist to account for hypothetical reasoning.
[63]See Beall and Murzi (2013).
[64]See Ripley (2013).
[65]See for example Burgis and Bueno (2019).

virtue of PUS dialetheism should handle it in exactly the same way it handles the others. But the strategy to avoid the Curry's paradox is not the same and PUS is not satisfied. Thus, this second issue represents a local objection to dialetheism.[66]

1.5 Summing up

Thus, there are some reasons why we might want to endorse dialetheism. But there are also some reasons to resist it. And it is unclear whether it is the most rational choice to make: the needle of the scale is not clearly skewed either for or against dialetheism. A good strategy to help settle the matter is to look for other advantages and/or disadvantages that are related to such a view. This is what I will try to do in the next chapters.

[66] Beall (2014) admits this is crucial problem.

Chapter 2

The denial and the rejection paradoxes

2.1 Preliminary remarks

Pragmatics is the field where most of the substantial objections to dialetheism fall.[1] We have mentioned the deficiency of Priest's pragmatic solution to the exclusivity problem in §1.4.1. Here, I want to go deeper in this issue.[2] So, let's present the problem once again in a slightly different way.

In classical logic, negation is exclusive in the sense that α and $\neg\alpha$ are incompatible: they cannot be both true. On the contrary, as we have seen in §1.2.3, negation is not exclusive in Priest's dialetheic logic: α and $\neg\alpha$ can sometimes be both true. Consequently, all the notions in which negation is deeply embedded fail to be exclusive as well. Consider, for example, «*in*compatible». Since it corresponds to «*not* being compatible», given the switch to dialetheism we can have a situation where what is incompatible is also compatible. Further significant examples are those of «*only* true» and «*only* false». They are usually understood as «true and not false» and «false and not true», respectively. Thus, in a dialetheic account a sentence can be e.g. both only false and true – as for «This sentence is only false». Recalling what we said in §1.4.1, dialetheic semantics lacks the resources to express exclusivity and this has been criticized on the basis of the following line of reasoning:

(P_1) dialetheic logic cannot express exclusivity

(P_2) exclusivity is necessary

(C) dialetheic logic is not correct

We have already justified both (P_1) and (P_2) in §1.4.1. But the argument is valid only if we add another premise, namely:

[1]This is what Priest (2006b, p. 94) himself admits.

[2]The present chapter is a more extended version of Carrara, Mancini, and Strollo (2021).

(P_3) exclusivity can only be expressed at the logical level

Basically, (P_3) denies that it is possible to recapture exclusivity at the pragmatic level. But Priest's claim is exactly that this is actually possible. First, he points out that the dialetheist's position is no worse than that of the classical logician:

> [a] dialetheist can express the claim that something, α, is not true – in those very words, $\neg T\langle\alpha\rangle$.[3] What she cannot do is ensure that the words she utters behave consistently: even if $\neg T\langle\alpha\rangle$ holds, $\alpha \wedge \neg T\langle\alpha\rangle$ may yet hold. But in fact, a classical logician (or anyone else who subscribes to the validity of Explosion) can do no better. He can endorse $\neg T\langle\alpha\rangle$, but this does not prevent his endorsing α as well. Of course, if he does (and assuming the T-schema), he will be committed to everything. But classical logic, as such, is no guard against this
>
> Priest (2006b, p. 291)

Second, he accepts both (P_1) and (P_2), but rejects (P_3), so that (C) does not follow. For he takes acceptance/rejection and assertion/denial to be exclusive and argues that it is possible to rely only on such exclusive pragmatic notions to get back the way we normally reason and convey information. In this perspective, exclusivity would not be embedded in the language and its logic, but in the way we use them.

Against Priest, Shapiro (2004) points out that relying only on exclusive assertion and denial is not sufficient to account for hypothetical reasoning. He writes:

> But how would a dialetheist formulate a hypothesis that someone is mistaken? Suppose that Karl says 'β', and his dialetheist friend Seymore does not wish to disagree (yet), but he wonders if Karl is mistaken. Seymore might want to assert a conditional in the form: 'if Karl is mistaken, then ϕ'. How can Seymore express this? Again, 'if $\neg\beta$ then ϕ' won't work. Since, for Seymore, $\neg\beta$ is compatible with β, it is not the way for him to say that Karl is mistaken in asserting β.
>
> Shapiro (2004, pp. 339-340)

That is, hypothetical reasoning requires us to make some exclusive assumptions (the antecedent of the conditional) without asserting or denying anything. The moral is that exclusivity does appear to be involved in different situations other than assertion and denial.

With this in mind, there are two possible strategies we can follow. The first one is, of course, to use this evidence against Priest's dialetheic semantics: what hypothetical reasoning shows is that exclusivity operates at the content level and, since dialetheic semantics is not capable of expressing exclusive notions, it is just

[3] Here the angle brackets are used as an appropriate name-forming device.

wrong. Alternatively, we could try to save dialetheic semantics by resorting to a further exclusive primitive act (and its related attitude): exclusive assumption. Granted – a dialetheist could say –, when we think hypothetically we do not assert and neither deny our hypothesis; nevertheless, we do perform a speech act: that of assuming (the hypothesis). And this act can be performed in an exclusive way, that is we can assume hypothesis as *only* true or *only* false (in addition to *at least* true and *at least* false).

In the next sections of this chapter, we will follow this strategy and develop a deductive system – DLEAC, the logic originally conceived by Carrara and Martino (2019) – which realizes it. Unfortunately, this approach will not prove sufficient to escape all problems. As I am going to show, or at least to argue, there is a further reason why the pragmatic recapture of exclusivity is not satisfying: revenge. In this regard, I will present and discuss two paradoxes – the *denial paradox* and *the rejection paradox*– that, I claim, represent a form of revenge and a threat for dialetheism. For their formal introduction, I will make use of a slightly extended version of DLEAC.

2.2 DLEAC

DLEAC – Dialetheic Logic with Exclusive Assumptions and Conclusions – is a dialetheic logic whose aim is to recapture exclusivity by modeling the speech acts of assuming and concluding.

According to Priest, acceptance and rejection are exclusive, and LEM holds in his dialetheic logic. Thus, it is possible for him to assert both α and $\neg\alpha$, but it is not possible to reject both of them: if we could do that, we would violate LEM. Therefore, that the subject S rejects α expresses that S considers α as *only* false. Conversely, that the subject S rejects $\neg\alpha$ expresses that S considers α as *only* true. Now, Carrara and Martino (2019) note that «[t]his use of rejection suggests the idea of a theory of natural deduction, where the acts of assuming and concluding may be understood in an ordinary or in an exclusive mode». Arguably, from a dialetheic point of view we can assume α in two different way: as true and false (inconsistently), or as only true (consistently). Let us call the former the *ordinary assumption* and the latter the *exclusive assumption*. From that, we can also model the act of concluding: to *ordinarily conclude* that α (under certain assumptions) means to prove that α is *at least* true, whereas to *exclusively conclude* that α (under certain assumptions) means to prove that α is *only* true. Therefore, the acts of concluding α and $\neg\alpha$ in an exclusive mode are incompatible in the sense that together lead to the rejection of some assumptions they depend on (a dialetheic form of the *reductio ad absurdum*). Thus, we can develop a logic – DLEAC – that formalises such notions.

Before presenting DLEAC's semantics and deductive system, some clarifications on exclusive/ordinary assumptions and conclusions are required. First, the act of concluding is modelled upon that of assuming. Thus, we can think of assumption as a primitive notion, whereas conclusion as being defined on it (plus the logic we choose). Second, it is better to think of assumption as a propositional attitude. In short: a subject S exclusively assumes α iff S's cognitive state is such that α is considered as *only* true; instead, a subject S ordinarily assumes α iff S's cognitive state is such that α is considered as *at least* true.[4] Then, the speech acts of exclusive and ordinary assumptions are the linguistic expressions of such attitudes. Third, as far as I know Priest is silent about exclusive/ordinary assumptions. Nonetheless, as we have previously discussed, it seems unavoidable for a dialetheist to appeal to such propositional attitudes in order to account for hypothetical reasoning. In conclusion, DLEAC can arguably be considered a legitimate dialetheic logic in line with Priest's perspective.

2.2.1 DLEAC: semantics

DLEAC's semantics is devised upon that of the logic of paradox (LP), i.e. the extensional part of the dialetheic logic we have introduced in chapter 1. Because of that, I refer to §1.2.2 for its discussion and here I will focus only on what differs from LP. For the sake of simplicity we can just omit function symbols. In addition, we will say that a sentence α is true (or at least true) if $1 \in v(\alpha)$, is false (or at least false) if $0 \in v(\alpha)$, is exclusively true if $0 \notin v(\alpha)$ and is exclusively false if $1 \notin v(\alpha)$.

Then, LP's semantics is extended by introducing a different notion of model. For this is crucial to model the acts (attitudes) of exclusive/ordinary assumptions. Let S be any set of sentences of our formal language, $\mathcal{L}$, some of which may be starred (i.e. marked by a star, $*$). Note that the star symbol $*$ does not belong to $\mathcal{L}$ and it can be thought of as a pragmatic operator: the exclusive assumption pragmatic operator. Thus, α^* means that α is considered as exclusively true. Therefore:

> A model M of S is an LP-interpretation in which all sentences of S are true and the starred ones are exclusively true.

Consequently, we have also a new notion of semantic consequence:

> A sentence α (a starred sentence α^*) is a semantic consequence of a set S of possibly starred sentences, $S \models \alpha^{(*)}$, if it is true (exclusively true) in every model of S.

The symbol $(*)$ is meant to provide a double reading of the formal expressions in which it occurs. For example, $S \models \alpha^{(*)}$ stands for $S \models \alpha$ and $S \models \alpha^*$. Moreover, when

[4] Note that this is not a definition, since we take assumption to be a primitive notion. It is just an helpful description, in the same manner as «a collection of elements which are considered as just one thing» (Cantor, 1882) is for the primitive notion of set.

the star refers to a complex formula, e.g. $\alpha \wedge \beta$, we will use the following notation where the underbraces and the overbraces show the scope of the star:

$$\underbrace{\alpha \wedge \beta}_{(*)} \quad \text{or} \quad \overbrace{\alpha \wedge \beta}^{(*)}$$

2.2.2 DLEAC: proof-theoretic system

Let us now introduce a proof system for DLEAC. Let $\alpha, \beta, \gamma, \ldots$ formulas of $\mathcal{L}$ and let $\Gamma, \Delta, \Lambda, \ldots$ sets of such formulas – as always, these sets can even be empty. Then, let an expression of the form $\Gamma \vdash C^{(*)}$ be a sequent, to be read «from the assumptions in Γ, one can infer the conclusion C (in an ordinary or exclusive mode)». The basic inferential rules of the proof system are shown below. When some stars occur in parentheses, $(*)$, the rule holds in the double form: (1) with all stars in parentheses at work and (2) with all stars in parentheses removed.

Reflexivity (R)

$$\alpha^{(*)} \vdash \alpha^{(*)}$$

$$\alpha^{*} \vdash \alpha$$

Weakening (W)

$$\frac{\Gamma \vdash \alpha^{(*)}}{\Gamma, \Delta \vdash \alpha^{(*)}}$$

Cut

$$\frac{\Gamma \vdash \alpha^{(*)} \qquad \Delta, \alpha^{(*)} \vdash \beta}{\Gamma, \Delta \vdash \beta}$$

$$\frac{\Gamma \vdash \alpha^{(*)} \qquad \Delta, \alpha^{(*)} \vdash \beta^{*}}{\Gamma, \Delta \vdash \beta^{*}}$$

Introduction of Absurd (IA)

$$\frac{\Gamma \vdash \alpha^{*} \qquad \Delta \vdash \neg\alpha}{\Gamma, \Delta \vdash \underbrace{\alpha \wedge \neg\alpha}_{*}}$$

Double Negation (DN)

$$\alpha^{(*)} \dashv\vdash \neg\neg\alpha^{(*)}$$

Disjunction

$$(\mathrm{I}\vee)\ \frac{\Gamma \vdash \alpha^{(*)}}{\Gamma \vdash \underbrace{\alpha \vee \beta}_{(*)}}$$

$$(\mathrm{E}\vee)\ \frac{\Gamma, \alpha \vdash \gamma^{(*)} \qquad \Delta, \beta \vdash \gamma^{(*)} \qquad \Lambda \vdash \alpha \vee \beta}{\Gamma, \Delta, \Lambda \vdash \gamma^{(*)}}$$

$$(\mathrm{E}\vee)\ \frac{\Gamma, \alpha^{*} \vdash \gamma^{(*)} \qquad \Delta, \beta^{*} \vdash \gamma^{(*)} \qquad \Lambda \vdash \overbrace{\alpha \vee \beta}^{*}}{\Gamma, \Delta, \Lambda \vdash \gamma^{(*)}}$$

Conjunction

$$(\mathrm{I}\wedge)\ \frac{\Gamma \vdash \alpha^{(*)} \qquad \Delta \vdash \beta^{(*)}}{\Gamma, \Delta \vdash \underbrace{\alpha \wedge \beta}_{(*)}}$$

$$(\mathrm{E}\wedge)\ \frac{\Gamma \vdash \overbrace{\alpha \wedge \beta}^{(*)}}{\Gamma \vdash \alpha^{(*)}}$$

$$(\mathrm{E}\wedge)\ \frac{\Gamma \vdash \overbrace{\alpha \wedge \beta}^{(*)}}{\Gamma \vdash \beta^{(*)}}$$

Identity

$$(\mathrm{I}=)\quad \vdash t = t$$

$$(\mathrm{E}=)\quad x = y, Px \vdash Py$$

$$(\mathrm{E}=)\ \frac{\Gamma, \alpha^{*} \vdash \overbrace{\neg(t = t)}^{*}}{\Gamma \vdash \neg\alpha}$$

$$(\mathrm{E}=)\ \frac{\Gamma, \alpha \vdash \overbrace{\neg(t = t)}^{*}}{\Gamma \vdash \underbrace{\neg\alpha}_{*}}$$

***Reductio ad Absurdum* (RAA)**

$$\frac{\Gamma, \alpha^{*} \vdash \overbrace{\beta \wedge \neg\beta}^{*}}{\Gamma \vdash \neg\alpha}$$

$$\frac{\Gamma, \alpha \vdash \overbrace{\beta \wedge \neg\beta}^{*}}{\Gamma \vdash \underbrace{\neg\alpha}_{*}}$$

As usual, the rules for the quantifiers are analogous to those of conjunction and disjunction.

The informal reading of DLEAC's fundamental rules, as well as some quick comments on them, might be helpful:

(R): from the assumption that α is at least true (exclusively true) we can infer the very same thing. Besides, from the assumption that α is exclusively true we can conclude that α is at least true.

(W): if we can (ordinarily or exclusively) conclude α from some premises, the assumption of additional premises does not alter the validity of the inference.

(Cut): if we can ordinarily (exclusively) conclude α from the premises in Γ, and if we can ordinarily deduce β from the premises in Δ plus the ordinary (exclusive) assumption of α, then we can ordinarily infer β from the premises in Γ and Δ. Besides, the same applies in case we can conclude β in an exclusive way from the premises in Δ plus the ordinary (exclusive) assumption of α, except that now we deduce β exclusively.

(I$\wedge$): if we can ordinarily (exclusively) conclude α from the premises in Γ, and if we can ordinarily (exclusively) conclude β from the premises in Δ, then we can ordinarily (exclusively) infer the conjunction $\alpha \wedge \beta$ from the premises in Γ plus those in Δ.

(E$\wedge$): if we can ordinarily (exclusively) infer a conjunction from some premises, then we can we can ordinarily (exclusively) infer each of its conjuncts from the same premises.

(I$\vee$): if we can ordinarily (exclusively) conclude α from some premises, then we can ordinarily (exclusively) conclude the disjunction $\alpha \vee \beta$, for any β.

(E$\vee$): if we can ordinarily (exclusively) infer γ both from the ordinary assumption of α plus the premises in Γ and from the ordinary assumption of β plus the premises in Δ, and if we can ordinarily conclude $\alpha \vee \beta$ from another set of premises, Λ, then we can ordinarily (exclusively) infer γ from the premises in Γ plus those in Δ and Λ. Besides, if we can ordinarily (exclusively) infer γ both from the exclusive assumption of α plus the premises in Γ and from the exclusive assumption of β plus the premises in Δ, and if we can exclusively conclude $\alpha \vee \beta$ from another set of premises, Λ, then we can ordinarily (exclusively) infer γ from the premises in Γ plus those in Δ and Λ.

(DN): from the ordinary (exclusive) assumption of α we can ordinarily (exclusively) conclude $\neg\neg\alpha$, and *vice versa*.

(IA): if we can exclusively infer α from the premises in Γ, and if we can ordinarily infer $\neg\alpha$ from the premises in Δ, then we can exclusively conclude $\alpha \wedge \neg\alpha$, i.e. a genuine absurd. As informal justification for (IA) we can think as follows. From α and $\neg\alpha$ we get $\alpha \wedge \neg\alpha$ through (I$\wedge$). Since α is only true it is not a dialetheia, and the same holds for $\neg\alpha$. Therefore, neither $\alpha \wedge \neg\alpha$ can be a dialetheia; thus, it is only true. Now, given that $\neg(\alpha \wedge \neg\alpha)$ is a dialetheic logical truth, the conclusion $\alpha \wedge \neg\alpha$ is authentically absurd – i.e it cannot be accepted even by a dialetheist.

(RAA): if we can conclude a genuine absurd from some premises and the exclusive assumption of α, then we can ordinarily infer $\neg\alpha$. Moreover, if we can conclude a genuine absurd from some premises and the ordinary assumption of α, then we can exclusively infer $\neg\alpha$.

(I=): For every object in the domain, it is a theorem that the object is self-identical. Note that here «t» is meant to represent either a variable or a constant.

(E=): First, the law of the indiscernibility of identical[5] holds: if we know that two objects are identical and that one of them is P, then also the other object is P. Second, $\underbrace{\neg(t = t)}_{*}$ is authentically absurd: if $\neg(t = t)$ is true, it must be a dialetheia.

In addition, here are some useful derived rules for DLEAC. Their proofs are shown in the appendix 2.5.

Material conditional
$$\frac{\Gamma, \alpha^{(*)} \vdash \beta^{(*)}}{\Gamma \vdash \neg\alpha \vee \beta}$$ $$\frac{\Gamma, \alpha^{*} \vdash \beta}{\Gamma \vdash \neg\alpha \vee \beta}$$ $$\frac{\Gamma, \alpha \vdash \beta^{*}}{\Gamma \vdash \underbrace{\neg\alpha} _{*} \vee \beta}$$

Modus Ponens (MPP)
$$\frac{\Gamma \vdash \alpha^{*} \qquad \Delta \vdash \neg\alpha \vee \beta}{\Gamma, \Delta \vdash \beta}$$ $$\frac{\Gamma \vdash \alpha \qquad \Delta \vdash \overbrace{\neg\alpha \vee \beta}^{*}}{\Gamma, \Delta \vdash \beta^{*}}$$

Law of Non-Contradiction (LNC)
$\Gamma \vdash \neg(\alpha \wedge \neg\alpha)$

Law of Excluded Middle (LEM)
$\Gamma \vdash \alpha \vee \neg\alpha$

Elimination of Absurd (EA)
$$\frac{\Gamma \vdash \overbrace{\alpha \wedge \neg\alpha}^{*}}{\Gamma \vdash \beta^{*}}$$

[5]This law is normally written as $\forall x \forall y(x = y \rightarrow \forall F(Fx \leftrightarrow Fy))$ and represents the left-to-right direction of the Leibniz's Law – whereas its right-to-left direction is the law of the identity of indiscernibles: $\forall x \forall y(\forall F(Fx \leftrightarrow Fy) \rightarrow x = y)$.

De Morgan			
$\dfrac{\Gamma \vdash \overbrace{\neg(\alpha \land \beta)}^{(*)}}{\Gamma \vdash \underbrace{\neg\alpha \lor \neg\beta}_{(*)}}$	$\dfrac{\Gamma \vdash \overbrace{\neg\alpha \lor \neg\beta}^{(*)}}{\Gamma \vdash \underbrace{\neg(\alpha \land \beta)}_{(*)}}$	$\dfrac{\Gamma \vdash \overbrace{\neg(\alpha \lor \beta)}^{(*)}}{\Gamma \vdash \underbrace{\neg\alpha \land \neg\beta}_{(*)}}$	$\dfrac{\Gamma \vdash \overbrace{\neg\alpha \land \neg\beta}^{(*)}}{\Gamma \vdash \underbrace{\neg(\alpha \lor \beta)}_{(*)}}$

As for LP, the material implication does not model a genuine conditional since MPP is not generally valid. Nevertheless, DLEAC validates MPP whenever at least one of the two premises (i.e. α and $\neg\alpha \lor \beta$) is only true. Note also that DLEAC validates a dialetheic version of the principle of Explosion, i.e. EA.

Finally, the Tarski's rules for the truth predicate, T. Here, as in chapter 1, we assume that $\underline{\alpha}$ is the name of α and that a unique name for every sentence is included into the domain.

T-scheme
$\dfrac{\Gamma \vdash \alpha^{(*)}}{\Gamma \vdash \underbrace{T\underline{\alpha}}_{(*)}}$
$\dfrac{\Gamma \vdash \overbrace{T\underline{\alpha}}^{(*)}}{\Gamma \vdash \alpha^{(*)}}$

T-scheme derived rules
$\dfrac{\Gamma \vdash \overbrace{\neg T\underline{\alpha}}^{(*)}}{\Gamma \vdash \underbrace{T\underline{\neg\alpha}}_{(*)}}$
$\dfrac{\Gamma \vdash \overbrace{T\underline{\neg\alpha}}^{(*)}}{\Gamma \vdash \underbrace{\neg T\underline{\alpha}}_{(*)}}$

2.3 The denial paradox

Let us focus again on assertion and denial. As should now be clear, the dialetheic denial of α is stronger than the assertion of $\neg\alpha$: allowing for the possibility of asserting both α and $\neg\alpha$, the *denial equivalence* must be rejected on pain of the lack of exclusivity. However, as Littmann and Simmons (2004) point out, this kind of relation between assertion and denial is completely non-standard. Because of that, the dialetheist is required to give a satisfactory account of such pragmatic notions. And, for such an account to be adequate, it must prove to be paradox-free – better, it must prove paradox-free with respect to the paradoxes that are unacceptable even for the dialetheist.

Priest (2006a, §6.4) claims that his notion of denial cannot yield to any relevant contradiction. In support of this, he discusses three cases which are supposed to cover all the possible situations that could be responsible for some paradox to come up. Let $\dashv$ be the force operator for denial, so that $\dashv \alpha$ represents the denial of α. As such, $\dashv \alpha$ is not a sentence but a speech act. Now, let us consider the following three examples:

(u) $\dashv u$ is true

(v) $\dashv$ the content of v is true

(w) $\dashv$ the content of w is false

Since $\dashv u$ is an utterance, it is not the kind of thing that can be true or false. Therefore, (u) does not make any sense and yields to no contradiction. About the utterance (v), its content is «the content of v is true» and what follows from the denial of it depends on whether we consider the content to be a truth-value gap or a truth-value glut. If we take it to be neither true nor false, then to deny (v) is the right thing to do; if we take it to be both true and false, we should not deny it since is true. In any case, (v) does not cause any contradiction. Finally, the utterance (w). Its content is «the content of w is false», i.e. a common liar sentence. If we take it to be glutty, we should assert it: its denial is just the wrong speech act we can opt for. If we take it to be gappy, then to deny it is the right speech act to perform. Thus, even (w) does not cause any contradiction. To sum up, with Priest's words:

> [w]hat these examples illustrate is the fact that attempts to formulate distinctive Liar paradoxes in terms of denial fail, since $\dashv$, being a force-operator, has no interaction with the content of what is uttered.
>
> Priest (2006a, p.108)

This, I claim, is just too fast and Priest's optimism must be tested further. For he does not seem to consider a very simple strategy through which the spectrum of paradoxes may cast its shadow again. Instead of using force operators, we can introduce the pragmatic notions as predicates that refer to the very same sentence in which they occur.

2.3.1 The (deceptive) assertion paradox

The strategy I have just mentioned is exactly that applied by Littmann and Simmons (2004) in their *assertion paradox*. Consider the sentence (α):

(α) α is not assertible

The authors defend that (α) is a dialetheia with the following informal proof:

Informal proof

Case 1) Suppose that α is true. Then, it follows that:

(1) it is the case of what it says;

(2) it is assertible, since it is true.

From (1) we can derive that α is not assertible, in contradiction with (2). Thus, α is both assertible and not assertible. But then, α is both true and false. Therefore, α is a dialetheia.

Case 2) Suppose that α is false. Then, it follows that:

(3) it is not the case of what it says;

(4) it is not assertible, since it is false.

From (3), α is assertible, in contradiction with (4). Thus, α is both assertible and not assertible. But then, α is both true and false, i.e. α is a dialetheia.

From the Excluded Middle we conclude that α is a dialetheia.

■

Now, if α is a dialetheia it is both assertible and not assertible. It is assertible because we have a proof that supports the truth of α; it is not assertible because this is exactly what it veridically says. But this would be something unacceptable even for the dialetheist. Thus, the assertion paradox would seem to threaten the legitimacy of dialetheism.

The proof given by Littmann and Simmons (2004) suffers from some problems, however. To begin with, the steps (2) and (4) are not correct. Let us start with the former. The mere supposition that α is true does not imply that α is also assertible. As we have seen in §1.2.3, for α to be assertible we need some good evidence in support of its truth, and a simple supposition does not provide such evidence. For the same reason, (4) is also incorrect: the supposition that α is false is not sufficient for not asserting it. Moreover, there is a further explanation why (4) is wrong. For a dialetheist, a sentence can be assertible even if there is good evidence that is false. For such a sentence might be a dialetheia, so that it is also true (as well as false). Thus, if we can prove that this sentence is a dialetheia, we can legitimately assert it. Finally, there is also a problem with the contradictory conclusion of the whole argument: α is both assertible and not assertible. Is this conclusion really problematic for a dialetheist? The answer is negative. Given that negation is not exclusive, «being assertible» and «not being assertible» do not exclude each other. (Instead, what she cannot allow is that there are sentences both assertible and deniable.) Thus, it would be possible for a dialetheist to accept sentences both assertible and not assertible – even if it is not at all clear what this would mean. In conclusion, the argument offered by Littmann and Simmons (2004) is not valid: dialetheism seems to be safe. But, is that so? In what follows I develop the Littman and Simmons' intuition in a different and simple way. The idea is to produce a paradox using the predicate of denial, i.e. «being deniable», and without resorting to negation. I will present such a *denial paradox* both informally and formally. For this latter purpose, I will extend DLEAC to obtain a logic – D-DLEAC – capable of expressing the predicate of denial.

2.3.2 The denial paradox: informal presentation

Let's start with the informal presentation of the paradox. Consider the following sentence:

$$(\delta) \qquad \delta \text{ is deniable}$$

I now intend to prove that δ is both assertible and deniable.

Informal proof

Let assume that δ is true. Then, it is deniable in virtue of what it says. Therefore, there is a state of knowledge for an arbitrary (rational) subject S such that S rejects δ and can (consequently) deny it. In the act of denying δ, S recognizes it as true. Such a recognition counts as good evidence in support of the truth of δ and, consequently, S can assert it. Thus, δ is assertable. Therefore, δ is both assertable and deniable. But (according to Priest's pragmatic account) this conclusion must be rejected. Then, δ cannot be true. Now, since the reasoning we have developed so far is a good evidence in support of the untruth of δ, S can deny it. But then, in the act of denying δ, S recognizes that what δ says is true. Therefore, given this evidence S can assert δ, so that it is both assertable and deniable.

■

This informal proof certainly needs some clarifications. There are (at least) three issues we need to discuss. First, for this argument to work we have to assume that (i) S is able to recognize the truth of δ as a result of the act of denying it, and that (ii) S is capable of making valid inferences and draw conclusions such those involved in this proof. More precisely, with respect to (i) S is required to accomplish the following tasks: to understand the meaning of δ, to grasp the notions of truth and untruth, and to know how to perform the acts of assertion and denial. All this – I claim – is not so much after all, and I take whatever rational subject to satisfy such requirements. Second, it could be argued that S's recognition of δ's truth in the act of denying it does not count as good evidence. Against this claim I reply by referring to Priest (2006b, p. 101-102),[6] where he accepts that direct experience (as in the case under discussion) can be legitimately considered good evidence. As a further clarification, let's think to this very simple situation. You and I are in a room with a unique closed window. I say: «Outside here it's raining». Now, the easiest way for you to check whether this sentence is true is to open the window and look outside. If your direct experience confirm what I've said, then you also have good evidence to assert it. And – I submit – this circumstance is analogous to the one where S recognises the truth of δ by denying it. Third, in this informal proof assertion and denial are based on the notions of acceptance and rejection described in §1.2.3, where the latter is disciplined by the condition **Reject(U)**: one ought to reject something if there is good evidence for its untruth. Thus, a possible way out from the paradox is to go for the already mentioned different version of this norm, i.e. **Reject$_2$(U)**: one ought to reject something if there is good evidence for its untruth, unless there is also good evidence for its truth. I postpone the analysis of this very issue to §2.4. Now, let's see how to formalise the paradox.

2.3.3 The denial paradox: formal presentation

To make things formal, we need to extend DLEAC (both its language and deductive system) to get a slightly richer logic we can call D-DLEAC. Let's start including the predicate D – to be read as «being deniable» – in $\mathcal{L}$. With that, we can now formalise δ:

$$(\delta) \quad D\underline{\delta}$$

Next, we incorporate the following basic rules in the deductive system.

[6] This passage has been already quoted in §1.2.3.

Rules for the predicate of denial	
(ED) $\dfrac{\Gamma \vdash D\underline{\alpha}}{\Gamma \vdash \underbrace{\neg\alpha}_{*}}$	(ID) $\dfrac{\Gamma \vdash \overbrace{\neg\alpha}^{*}}{\Gamma \vdash D\underline{\alpha}}$

These rules are designed to regiment D's logical behaviour as described by Priest, and some comments are necessary. On the one hand, I believe that ED shouldn't be controversial. It states that if α is deniable, then we can conclude that it is exclusively false (i.e. its negation is exclusively true). In other words, ED expresses the idea that for a sentence to be exclusively false is a necessary condition for being deniable, and this is in line with Priest's view. On the other hand, ID states that if α is exclusively false, then we can infer that it is deniable, and this can be questioned. For it may be objected that for a sentence to be exclusively false is not sufficient for being deniable: we need also good evidence in support of it, as well as a subject who is able to perform such an act. My reply is that I no doubt agree with this general criticism,[7] but I argue that it does not affect our specific situation: the denial paradox. To see why, I now show and discuss its proof.

Proof

1	(1)	$D\underline{\delta}$	Assumption
1	(2)	$\underbrace{\neg D\underline{\delta}}_{*}$	1 ED
1	(3)	$\underbrace{D\underline{\delta} \wedge \neg D\underline{\delta}}_{*}$	1, 2 IA
	(4)	$\underbrace{\neg D\underline{\delta}}_{*}$	1, 3 RAA
	(5)	$D\underline{\delta}$	4 ID
	(6)	$\underbrace{D\underline{\delta} \wedge \neg D\underline{\delta}}_{*}$	4, 5 IA

[7] As one of the referees pointed out, it is important to emphasize that this concern is serious, and is a good reason to reject ID, making D not well-defined. In other words, the proposed logic fails to give a proper formalization of the denial predicate in Priest's dialetheic view. However, my claim is that it is a safe and good enough tool I can legitimately use for my purpose.

With respect to the criticism about $\mathsf{I}D$, the suspicions might fall on the step (5). But I argue that (5) is not 'guilty of any crime'. The reason is that, in this case, we do have good evidence in support of $\underbrace{\neg D\underline{\delta}}_{*}$, so that we can legitimately conclude $D\underline{\delta}$. Such an evidence is represented by the inference (1)-(4), which is part of the proof. Therefore, that $D\underline{\delta}$ is exclusively false is well justified and the criticism does not apply here. Moreover, I also take the existence of a subject who can deny δ to be guaranteed. For there are arguably at least two subjects who meet this requirement: you (the one who read the proof) and I (the one who wrote it). Thus, step (5) is 'clean', the proof is valid and the result to which it leads is a paradoxical conclusion that is incompatible with Priest's pragmatic account. In addition, note that the paradox holds even for the classical logician who accepts that assertion/denial and acceptance/rejection are exclusive, and that $\mathsf{I}D$ and $\mathsf{E}D$ are correctly employed in the proof. The only difference for her is that the star is no longer necessary.

Moral of the story: since D-DLEAC faithfully reflects Priest's dialetheic logical machinery and since it allows to infer a conclusion he explicitly rejects, the paradox represents a threat for his dialetheic view. How then might Priest reply?

2.4 Ways out?

A first available solution to the denial paradox is the usual tarskian strategy: a restriction of the expressive capacity of $\mathcal{L}$ through the banishment of D, or at least by imposing that D cannot apply to the very same sentence where it occurs. However, this is hardly something Priest could be willing to accept for the same reasons he rejects the tarskian solution to the liar paradox (cf. §1.3.1). In this respect, consider that such pragmatic predicates like «deniable» and «rejectable» are commonly employed in our natural language and even Priest makes use of them.[8]

Another intriguing move to resist the paradox might be S's silence. Since – in the first part of the argument – for S to recognize δ as true she has to perform the act of denying δ, we can just 'hush her': being S silent, she lacks the evidence in support of δ's truth and then she cannot assert it. However, I find this alleged solution unsatisfactory for two reasons. First, if we take such a proposal seriously we get an inscrutable situation where a sentence is deniable (in principle) but cannot be (practically) denied, on pain of a dialetheically unacceptable contradiction. But what is deniability if not the practical possibility to deny something? Can we still say of a sentence which is deniable in principle but not practically deniable that it is deniable? Maybe, it might be suggested to clarify this confusing situation by distinguishing between the predicates «*practically* deniable» and «*theoretically* deniable».

[8]See for example Priest (2006b, p. 114): «from the facts that a disjunction is rationally acceptable and that one of its disjuncts is rationally both acceptable and rejectable, it does not follow that the other is rationally acceptable». Or again Priest (2006b, p. 22): «it would seem most implausible that the liar sentence should be assertible on Mondays, Wednesdays, and Fridays and deniable on Tuesdays».

But this approach does not solve the problem since we can just take the sentence «This sentence is practically deniable» and develop a similar paradoxical argument as before. Second, there is a stronger reason why this solution does not work: the *rejection paradox*. Let's just say, for the sake of argument, that S's silence prevents the paradoxical conclusion. Thus, we can rephrase the paradox in terms of rejection instead of denial, so that no act needs to be performed. Consider the sentence:

$$(\rho) \quad \rho \text{ is rejectable}$$

Then we can prove that ρ is both acceptable and rejectable.

Informal proof

Let assume that ρ is true. Then, it is rejectable in virtue of what it says. Therefore, there is a state of knowledge for an arbitrary (rational) subject S such that S rejects ρ. Experiencing such a cognitive state, S recognizes ρ as true. Such a recognition counts as good evidence in support of the truth of ρ and, consequently, S can accept it. Thus, ρ is acceptable. Therefore, ρ is both acceptable and rejectable. But (according to Priest's pragmatic account) this conclusion must be rejected. Then, ρ cannot be true. Now, since the reasoning we have developed so far is a good evidence in support of the untruth of ρ, S can reject it. But then, through the rejection of ρ, S recognizes that what it says is true. Therefore, given this evidence S can accept ρ, so that it is both acceptable and rejectable.

■

Note that the reason why the rejection paradox does not suffer from the alleged solution of S's silence is that an arbitrary rational subject cannot but reject (accept) something provided she has good evidence in support of its untruth (truth). In short, S cannot choose freely whether to believe or reject something: she is forced towards belief or rejection by her rationality.

Of course, the rejection paradox can be targeted by a similar criticism to the one related to S's requirements we have discussed in §2.3.2. For in the argument it is assumed that (i) S is able to recognize the truth of ρ as a result of its rejection, and that (ii) S is capable of making valid inferences and draw conclusions such those involved in this proof. More precisely, with respect to (i) S is required to accomplish the following tasks: to understand the meaning of ρ, to grasp the notions of truth and untruth, and to experience the attitudes of acceptance and rejection. Thus, my reply is analogous to the one of before: I take whatever rational subject to satisfy such requirements.

Given what we said in §2.3.3, the formal presentation of the rejection paradox should be quite straightforward. Let's extend DLEAC by including the predicate

R – to be read as «being rejectable» – in $\mathcal{L}$. Then, the logic we get can be named R-DLEAC. With R, we can formalise ρ:

$$(\rho) \quad R\underline{\rho}$$

Next, we add the following basic rules to the deductive system.

Rules for the predicate of rejection	
$(\mathsf{E}R)\ \dfrac{\Gamma \vdash R\underline{\alpha}}{\Gamma \vdash \underbrace{\neg\alpha}_{*}}$	$(\mathsf{I}R)\ \dfrac{\Gamma \vdash \overbrace{\neg\alpha}^{*}}{\Gamma \vdash R\underline{\alpha}}$

It should be evident that R-DLEAC and D-DLEAC are equivalent deductive systems: there is no difference between them, except for the symbol R instead of D, which of course is not relevant. Thus, being regimented by the same rules, D and R boil down to the very same thing. This is certainly wrong, since what they are supposed to capture, i.e. deniability and rejectability, are different predicates. The difference is that for something to be rejectable we need good evidence in support of its exclusive falsity, as well as a subject who experiences such a propositional attitude, instead of performing an act. Then, this deficiency is due (again) to the inadequacy of $\mathsf{I}R$ and $\mathsf{I}D$. But similar comments to those of §2.3.3 apply here, and then we don't need to worry about that: for our purposes we can reasonably use $\mathsf{I}R$. Thus, the proof obviously replicates that of the denial paradox:

Proof

1	(1)	$R\underline{\rho}$	Assumption
1	(2)	$\underbrace{\neg R\underline{\rho}}_{*}$	1 $\mathsf{E}R$
1	(3)	$\underbrace{R\underline{\rho} \wedge \neg R\underline{\rho}}_{*}$	1, 2 IA
	(4)	$\underbrace{\neg R\underline{\rho}}_{*}$	1, 3 RAA

(5)	$R\underline{\rho}$	4 IR
(6)	$\underbrace{R\underline{\rho} \wedge \neg R\underline{\rho}}_{*}$	4, 5 IA

Needless to say, the moral is the same of that of the denial paradox.[9]

All this considered, Priest seems to suggest the following two options which may help resisting the paradox: (i) a refinement of the pragmatical notions at stake, and (ii) the introduction of rational dilemmas. Let's start with (i). Priest already encountered a dialetheically unacceptable paradoxical situation, i.e. that caused by the strengthen liar sentence (i.e. «This sentence is untrue») and the following norms for acceptance and rejection:

> **Accept**. One ought to accept something if there is good evidence for its truth.
> **Reject(U)**. One ought to reject something if there is good evidence for its untruth.
>
> Priest (2006a, p. 109-110)

For since we do have good evidences in support of both its truth and untruth (the paradoxical proof, as it is usually developed), the strengthen liar sentence would be both acceptable/assertible and rejectable/deniable – the very same situation we get with the rejection/denial paradox. It is for this reason that Priest suggests to replace **Reject(U)** with the following condition:

> **Reject$_2$(U)**. One ought to reject something if there is good evidence for its untruth, unless there is also good evidence for its truth.
>
> Priest (2006a, p. 110)

With this move, the strengthen liar sentence becomes exclusively acceptable. Then, what about the rejection paradox? Let's try to develop the informal proof once again, assuming **Reject$_2$(U)** instead of **Reject(U)**.

Informal proof

Case 1)

[9] The reader may wonder why I haven't decided to present the rejection paradox only, since it is arguably stronger and maybe clearer then that of denial, and since it alone is enough to make emerge the problematic issue of a dialetheically unacceptable paradoxical sentence. The reason is that the way I have presented the matter here corresponds to the chain of thoughts I went trough in the process of thinking about this topic. Therefore, the kind of discussion I have opted for should be more informative for the reader.

Let assume that ρ is true. Then, it is rejectable in virtue of what it says. Therefore, there is a state of knowledge for a (rational) subject S such that S rejects ρ. Experiencing such a cognitive state, S recognizes ρ as true. Such a recognition counts as good evidence in support of the truth of ρ and, consequently, S can accept it. Thus, ρ is acceptable, and it is not rejectable anymore because of the condition **Reject$_2$(U)**. Therefore, ρ is false. In conclusion, ρ is an (only) acceptable dialetheia.

Case 2)
Let assume that ρ is false. Then, it is not the case of what it says. That is, it can be either acceptable or a third option which corresponds to neither acceptable nor rejectable.[10] Let assume that ρ is acceptable. Thus, there is a state of knowledge for an arbitrary (rational) subject S such that S accepts ρ. This implies that it is true and that there is good evidence in support of its truth. Therefore, ρ is an (only) acceptable dialetheia. On the other hand, if ρ is neither acceptable nor rejectable, S remains agnostic about it and no contradiction follows.

Finally, from Excluded Middle we can conclude that either ρ is an (only) acceptable dialetheia or an undecidable false sentence. Either ways, no dialetheically unacceptable contradiction comes up from ρ.

■

Reject$_2$(U) seems to save the day.[11] However, it is a very controversial condition and can be questioned on the basis of the unjustified asymmetry it involves. For **Reject$_2$(U)** makes acceptance and rejection (as well as assertion and denial) deeply asymmetric, in the sense that good evidences in favour of the truth of a sentence prevail over good evidences in support of the untruth of the same sentence. This makes rejection a never conclusive attitude, since a future discovery of good evidence of the truth of the sentence tips definitively the scale in favour of its acceptance. Such a situation, I claim, contrasts with our ordinary practice of rational inquiry, where acceptance and rejection seem to be at least equally balanced, if not unbalanced to the advantage of the latter. To see why, take a sentence whose truth value can be empirically determined, e.g. «Right now, it's raining in Paris». The direct observation of Paris' weather conclusively answers the question about whether this sentence should be accepted or rejected, both in the case we have good evidence in support of its truth and in the case we have good evidence of its untruth. The same occurs in a more abstract domain as that of mathematics. Normally, when a mathematician correctly disproves something, she rejects it and does not keep on trying also to prove

[10]This case seems to be contemplated by Priest, for example when a subject lacks any evidence about a sentence and remains agnostic about it. It could also be the case of some undecidable sentences, the truth value of which cannot be known by a rational subject in principle.

[11]With regard to the denial paradox the conclusion is analogous and the argument is almost identical. Let's call *suspension* the linguistic expression of a neither acceptable nor rejectable sentence. Thus, assuming Reject$_2$(U) we get that δ is either an (only) assertible dialetheia or a suspendible false sentence.

it: she considers the matter closed. The same point is defended also in the following passage of Murzi and Carrara (2015):

> [**Reject$_2$(U)**] makes denial profoundly *unlike* assertion. Unlike assertion, any denial may later turn out to be incorrect, since any false sentence can in principle be discovered to be a glut. Thus, you can disagree with my assertion that $0 \neq 0$, and thus deny $0 \neq 0$. But, even if you can prove $0 = 0$, and hence disprove $0 \neq 0$, you can never be fully confident that your denial is correct: a proof of $0 \neq 0$ may always turn up. By contrast, if you have proved $0 = 0$ and thereby assert it, you can be fully confident that your assertion is correct. We find this asymmetry problematic: nothing in our practice of asserting and denying things, it seems to us, suggests that assertion can be indefeasible in a way that denial is not.
>
> Murzi and Carrara (2015, p. 114)

Priest may reply that this is so because consistency is generally taken for granted, but that this (consistency) is exactly what is at issue. That is, he may go for a revisionistic approach: our ordinary practice of rationality is wrong when applied to actually inconsistent domains and we should change it on the basis of what dialetheism suggests. But since we cannot know in advance whether or not the issue under discussion is consistent, we had better not to assume that it is and always look for good evidences for its truth, despite having good evidences for its untruth. This, I claim, is what rationality demands according to **Reject$_2$(U)**. However, so far almost no one has proceeded in this way in a rational inquiry. Presumably, the (macroscopic) empirical world is consistent: for a given *phenomenon*, there cannot be observations in support of both truth and untruth. And this would explain why we immediately exclude the truth of something when we have evidence of its untruth. But, arguably, in abstract domains consistency is much more debatable. And yet, even there the usual rational inquiry has always been consistent.

Indeed, it could even be argued that in the empirical sciences acceptance is weaker than rejection, at least when what is accepted or rejected is a theory. This is to endorse Popper's falsification principle, according to which a theory should be accepted as long as it is empirically verified and not falsified. Now, it can be pointed out that **Accept** and **Reject$_2$(U)** refer to single sentences, and that acceptance and rejection of a theory may require different norms. However, a theory can be conceived as a unique sentence: the conjunction of all axioms and the sentences belonging to their logical closure.[12] Besides, the same comment applies also to a single natural law. Consider the sentence «Superluminal motion is not possible». We have good reason in support of its truth – i.e. General Relativity – and therefore we accept it. But in case in the future we observed e.g. a superluminal neutrino, we would be forced to reject such a sentence and, what is more, to do it conclusively. To sum up: **Reject$_2$(U)** does not seem to fit the way we normally practice rationality – a practice that benefits

[12] Though, this would be an infinite sentence and this may be considered problematic.

from a long time of corroborative successes and that does not appear to have an urge for such a profound revision.

Moreover, **Reject$_2$(U)** is victim of the very same problem it tries to solve: the exclusivity problem. Since there is no language/metalanguage distinction, the semantics in use to express **Reject$_2$(U)** is the dialetheic semantics with the usual non-exclusive negation. But the norm can be rephrased as:

> **Reject$_2$(U)**. (there is good evidence for α's untruth) $\wedge$ $\neg$ (there is good evidence for α's truth) $\rightarrow$ (one ought to reject α).

Therefore, because of the presence of $\neg$, the norm may prove to be unable to ensure exclusivity. And to rely again on rejection to recover it leads to circularity. Now, admittedly **Reject$_2$(U)** is not a definition but a norm of rationality, and rejection does not need any definition inasmuch it is primitive. But a norm whose aim is to supervise and control something (e.g. an attitude or an act), cannot depend on the very same notion it aims at supervising and controlling. If it does so, it is (badly) circular and, consequently, useless. And since it is useless, we had better to leave it behind.

A different refinement of the conditions for acceptance and rejection might be obtained by operating on the notion of good evidence, on which these norms depend. Let's consider the following proposal:

> **Accept$_3$**. One ought to accept something if there is (good) evidence for its truth better than the evidence for its untruth.
> **Reject$_3$(U)**. One ought to reject something if there is (good) evidence for its untruth better than the evidence for its truth.

Following these norms, the exclusivity of acceptance and rejection relies on the possibility to always compare the evidences and establish a strict order among them. Thus, the following questions become important: do exist some *criteria* we can use to evaluate an evidence? What are they? Do they always allow us to strictly order the evidences? Now, there is no need to develop a whole account of good evidences to settle the matter. For I claim there are some clear examples of equally valuable evidences which show us that **Accept$_3$** and **Reject$_3$(U)** are not suitable for Priest. For consider the strengthen liar and its usual proof. Here, the evidence in support of its truth is arguably of identical value of that in support of its untruth. For these two evidences are of the very same nature: (good) arguments. Thus, these two norms cannot determine whether the strengthen liar sentence should be accepted of rejected. A further move might be to allow a third attitude in case of a draw between evidences: to remain agnostic. But recall that, being true, Priest is willing to accept this sentence. In conclusion: assuming **Accept**$_3$ and **Reject**$_3$**(U)**, most of the dialetheias Priest accepts become sentences to be agnostic about, and since this does not correspond to the way of his dialetheism, we can just discard this proposal.

Let's now focus on Priest's second option, (ii): the introduction of rational dilemmas. We have already introduced what a rational dilemma is in §1.2.3. A typical example of such a dilemma is Littmann irrationalist's paradox.[13] Consider the following sentence:

(χ) It is irrational to believe χ

Let I be «It is irrational (to bring it about) that» and let B be «You believe that». Thus, χ is formalised as:

$$(\chi) \quad IB\chi$$

Let O be the operator from standard deontic logic «It is obligatory that», and let's assume the following uncontroversial principle about rationality: $IB(\alpha \wedge IB\alpha)$. Thus, we get a rational dilemma:

Proof

(1)	$IB(\chi \wedge IB\chi)$	$\vdash IB(\alpha \wedge IB\alpha)$
(2)	$IB(\chi \wedge \chi)$	1 $\chi = IB\chi$
(3)	$IB\chi$	2 $\alpha \wedge \alpha \dashv\vdash \alpha$
(4)	χ	3 $\chi = IB\chi$
(5)	$O\neg B\chi$	3 $IB\alpha \vdash O\neg B\alpha$
(6)	$OB\chi$	4 if $\vdash \alpha$ then $\vdash OB\alpha$
(7)	$OB\chi \wedge O\neg B\chi$	5, 6 I$\wedge$

[13]See Priest (2006a, p. 111-112).

Our rationality seems to force us to do something impossible and violate $\Box\neg(B\alpha \wedge \neg B\alpha)$:[14] it asks to believe and not to believe the same sentence simultaneously. This – Priest claims – is not caused by our defective comprehension of rationality and the way we account for it, but is simply a fact of life. Rational dilemmas would stem from the system of norms for rationality in the same way as moral and legal dilemmas are generated by the systems of ethical and juridical norms. Thus, we may wonder: is the rejection/denial paradox another example of rational dilemma?

I claim it is possible to get a rational dilemma from ρ (and δ). Let J be «You reject that». Thus:

Proof[15]

1	(1)	$R\underline{\rho}$	Assumption
1	(2)	$\underbrace{\neg R\underline{\rho}}_{*}$	1 $\mathsf{E}R$
1	(3)	$\underbrace{R\underline{\rho} \wedge \neg R\underline{\rho}}_{*}$	$1, 2$ IA
	(4)	$\underbrace{\neg R\underline{\rho}}_{*}$	$1, 3$ RAA
	(5)	$R\underline{\rho}$	4 $\mathsf{I}R$
	(6)	$OB\rho$	5 if $\vdash \alpha$ then $\vdash OB\alpha$
	(7)	$OJ\rho$	4 if $\vdash \neg\alpha$ then $\vdash OJ\alpha$
	(8)	$OB\rho \wedge OJ\rho$	7 I$\wedge$

But, according to Priest, $\overbrace{\Box\neg(B\alpha \wedge J\alpha)}^{*}$. Therefore, we have a dilemma: rationality forces us to do something impossible.

We can even say something more general. Given the inferences I used in steps (7), i.e. if $\vdash \neg\alpha$ then $\vdash OJ\alpha$, and (6), i.e. if $\vdash \alpha$ then $\vdash OB\alpha$, for every provable contradiction we can get a dilemma. This, of course, is something Priest would definitely reject:

[14] Arguably, Priest would state this principle as $\overbrace{\Box\neg(B\alpha \wedge \neg B\alpha)}^{*}$. But lacking an appropriate logical machinery to express modality with exclusive assumption – i.e. an alethic modal extension of DLEAC – this formalisation is rather ambiguous. However, such an extension is quite straightforward using a possible world semantics.

[15] A short comment about the proof: the rules I used in steps (6) and (7) shall be justified by the norms **Accept** and **Reject(U)**.

since dialetheias are meant to be exclusively acceptable, not all paradoxes are allowed to produce a dilemma. This is the reason why he suggests to use **Reject$_2$(U)**, which would greatly reduce the number of rational dilemmas since the rule in (7) becomes: if ($\vdash \neg\alpha$ and not $\vdash \alpha$) then $\vdash OJ\alpha$. Thus, since we do have a proof of ρ (steps (1)-(5)), the whole proof stops at step (6) and the dilemma vanishes.

Though, a paradox and a dilemma are two different kind of things. Priest is explicit about rational dilemmas: we should bear them since they are «facts of life». But he seems to be more reluctant to accept the paradoxes that he considers – guess what – unacceptable (e.g. the rejection/denial paradox). After all, dialetheism was originally conceived as a way to cope with some paradoxical situations by showing that they are what they *prima facie* seem to be, and that do not pose any problem. However, if a dialetheist allow for some dialetheically unacceptable paradoxes, she has to admit that somewhere in her overall view there is a problem. That is: dialetheism would be just a partial solution in need to be fixed or completed.

Now, these paradoxes appear to be caused by the rules of introduction and elimination of the predicates R and D, which in turn are related to the norms for acceptance and rejection. For it is on the basis of **Accept** and **Reject$_2$(U)** that we developed IR, ER, ID and ED. Moreover, also the dilemmas that originate from such paradoxes are caused by the same norms: steps (6) and (7) use inferences strictly related to **Accept** and **Reject$_2$(U)**. This means that both rational dilemmas and such paradoxes share a common root: an unsatisfactory regimentation (likely due to a limited comprehension) of these attitudes. Thus, if we have to accept rational dilemmas, why not to accept also the rejection/denial paradox as «a fact of life»? What is the reason why we should try to avoid it? I suspect that if Priest acknowledged some dialetheically unacceptable paradoxes his enterprise would result severely damaged, to the extent that dialetheism may lose its status of rationally preferable theory since its explanatory power would be reduced (not being able to account for such problematic paradoxes). But this entire chapter aims precisely at establishing that the rejection and the denial paradoxes actually are examples of such dialetheically unacceptable paradoxes and that, consequently, pose a threat to dialetheism.

2.5 Appendix: some relevant proofs in DLEAC.

Here, I show the proofs of the DLEAC's derived rules we have introduced in §2.2.2, together with that of the liar paradox – i.e. the proof that the liar sentence is a dialetheia. Because of its better ease of reading, I use the natural deduction Lemmon-style.

2.5.1 Proofs of some DLEAC's derived rules

The following proofs make use of DLEAC's basic rules only. There are just two exceptions: the proof of MPP (where I use De Morgan) and that of EA (where I use MPP). Thus, to avoid possible concerns I show the proofs in a non-circular order, so that when I use a derived rule, this has already been proved.

Law of Excluded Middle (LEM)

$$\Gamma \vdash \alpha \vee \neg\alpha$$

Proof

1	(1)	$\overbrace{\neg(\alpha \vee \neg\alpha)}^{*}$	Assumption
2	(2)	α^{*}	Assumption
2	(3)	$\overbrace{\alpha \vee \neg\alpha}^{*}$	2 I∨
1, 2	(4)	$\underbrace{(\alpha \vee \neg\alpha) \wedge \neg(\alpha \vee \neg\alpha)}_{*}$	1, 3 I∧
1	(5)	$\neg\alpha$	2, 4 RAA
1	(6)	$\alpha \vee \neg\alpha$	5 I∨
1	(7)	$\neg\neg(\alpha \vee \neg\alpha)$	6 DN
1	(8)	$\underbrace{\neg(\alpha \vee \neg\alpha) \wedge \neg\neg(\alpha \vee \neg\alpha)}_{*}$	1, 7 IA
	(9)	$\neg\neg(\alpha \vee \neg\alpha)$	1, 8 RAA
	(10)	$\alpha \vee \neg\alpha$	9 DN

De Morgan

$$\frac{\Gamma \vdash \overbrace{\neg(\alpha \wedge \beta)}^{(*)}}{\Gamma \vdash \underbrace{\neg\alpha \vee \neg\beta}_{(*)}}$$

Proof (non-starred version)

1	(1)	$\neg(\alpha \wedge \beta)$	Assumption
2	(2)	$\overbrace{\neg(\neg\alpha \vee \neg\beta)}^{*}$	Assumption
3	(3)	$\neg\alpha$	Assumption
3	(4)	$\neg\alpha \vee \neg\beta$	3 I$\vee$
3	(5)	$\neg\neg(\neg\alpha \vee \neg\beta)$	4 DN
2, 3	(6)	$\underbrace{\neg(\neg\alpha \vee \neg\beta) \wedge \neg\neg(\neg\alpha \vee \neg\beta)}_{*}$	2, 5 IA
2	(7)	α^*	3, 6 RAA
8	(8)	$\neg\beta$	Assumption
8	(9)	$\neg\alpha \vee \neg\beta$	8 I$\vee$
8	(10)	$\neg\neg(\neg\alpha \vee \neg\beta)$	9 DN
2, 8	(11)	$\underbrace{\neg(\neg\alpha \vee \neg\beta) \wedge \neg\neg(\neg\alpha \vee \neg\beta)}_{*}$	2, 10 IA
2	(12)	β^*	8, 11 RAA
2	(13)	$\underbrace{\alpha \wedge \beta}_{*}$	7, 12 I$\wedge$
1, 2	(14)	$\underbrace{(\alpha \wedge \beta) \wedge \neg(\alpha \wedge \beta)}_{*}$	1, 13 IA
1	(15)	$\neg\neg(\neg\alpha \vee \neg\beta)$	2, 14 RAA
1	(16)	$\neg\alpha \vee \neg\beta$	15 DN

Proof (starred version)

1	(1)	$\overbrace{\neg(\alpha \wedge \beta)}^{*}$	Assumption
2	(2)	$\neg(\neg\alpha \vee \neg\beta)$	Assumption
3	(3)	$\underbrace{\neg\alpha}_{*}$	Assumption
3	(4)	$\underbrace{\neg\alpha \vee \neg\beta}_{*}$	3 I∨
2, 3	(5)	$\underbrace{(\neg\alpha \vee \neg\beta) \wedge \neg(\neg\alpha \vee \neg\beta)}_{*}$	2, 4 IA
2	(6)	α	3, 5 RAA
7	(7)	$\underbrace{\neg\beta}_{*}$	Assumption
7	(8)	$\underbrace{\neg\alpha \vee \neg\beta}_{*}$	7 I∨
2, 7	(9)	$\underbrace{(\neg\alpha \vee \neg\beta) \wedge \neg(\neg\alpha \vee \neg\beta)}_{*}$	2, 8 IA
2	(10)	β	7, 9 RAA
2	(11)	$\alpha \wedge \beta$	6, 10 I∧
1, 2	(12)	$\underbrace{(\alpha \wedge \beta) \wedge \neg(\alpha \wedge \beta)}_{*}$	1, 11 IA
1	(13)	$\underbrace{\neg\neg(\neg\alpha \vee \neg\beta)}_{*}$	2, 12 RAA
1	(14)	$\underbrace{\neg\alpha \vee \neg\beta}_{*}$	13 DN

$$\frac{\Gamma \vdash \overbrace{\neg\alpha \vee \neg\beta}^{(*)}}{\Gamma \vdash \underbrace{\neg(\alpha \wedge \beta)}_{(*)}}$$

Proof (non-starred version)

1	(1)	$\neg\alpha \vee \neg\beta$	Assumption
2	(2)	$\neg\alpha$	Assumption
3	(3)	$\underbrace{\alpha \wedge \beta}_{*}$	Assumption
3	(4)	α^*	3 E∧
2, 3	(5)	$\underbrace{\alpha \wedge \neg\alpha}_{*}$	2, 4 IA
2	(6)	$\neg(\alpha \wedge \beta)$	3, 5 RAA
7	(7)	$\neg\beta$	Assumption
8	(8)	$\underbrace{\alpha \wedge \beta}_{*}$	Assumption
8	(9)	β^*	8 E∧
7, 8	(10)	$\underbrace{\beta \wedge \neg\beta}_{*}$	7, 9 IA
7	(11)	$\neg(\alpha \wedge \beta)$	8, 10 RAA
1	(12)	$\neg(\alpha \wedge \beta)$	1, 2, 6, 7, 11 E∨

Proof (starred version)

1	(1)	$\overbrace{\neg\alpha \vee \neg\beta}^{*}$	Assumption
2	(2)	$\underbrace{\neg\alpha}_{*}$	Assumption
3	(3)	$\alpha \wedge \beta$	Assumption
3	(4)	α	3 E∧

2, 3	(5)	$\underbrace{\alpha \wedge \neg\alpha}_{*}$	2, 4 IA
2	(6)	$\underbrace{\neg(\alpha \wedge \beta)}_{*}$	3, 5 RAA
7	(7)	$\underbrace{\neg\beta}_{*}$	Assumption
8	(8)	$\alpha \wedge \beta$	Assumption
8	(9)	β	8 E$\wedge$
7, 8	(10)	$\underbrace{\beta \wedge \neg\beta}_{*}$	7, 9 IA
7	(11)	$\underbrace{\neg(\alpha \wedge \beta)}_{*}$	8, 10 RAA
1	(12)	$\underbrace{\neg(\alpha \wedge \beta)}_{*}$	1, 2, 6, 7, 11 E$\vee$

$$\frac{\Gamma \vdash \overbrace{\neg(\alpha \vee \beta)}^{(*)}}{\Gamma \vdash \underbrace{\neg\alpha \wedge \neg\beta}_{(*)}}$$

Proof (non-starred version)

1	(1)	$\neg(\alpha \vee \beta)$	Assumption
2	(2)	α^*	Assumption
2	(3)	$\underbrace{\alpha \vee \beta}_{*}$	2 I$\vee$
1, 2	(4)	$\underbrace{(\alpha \vee \beta) \wedge \neg(\alpha \vee \beta)}_{*}$	1, 3 IA
1	(5)	$\neg\alpha$	2, 4 RAA
6	(6)	β^*	Assumption
6	(7)	$\underbrace{\alpha \vee \beta}_{*}$	6 I$\vee$

1, 6	(8)	$\underbrace{(\alpha \vee \beta) \wedge \neg(\alpha \vee \beta)}_{*}$	1, 7 IA
1	(9)	$\neg\beta$	6, 8 RAA
1	(10)	$\neg\alpha \wedge \neg\beta$	5, 9 I$\wedge$

Proof (starred version)

1	(1)	$\overbrace{\neg(\alpha \vee \beta)}^{*}$	Assumption
2	(2)	α	Assumption
2	(3)	$\alpha \vee \beta$	2 I$\vee$
1, 2	(4)	$\underbrace{(\alpha \vee \beta) \wedge \neg(\alpha \vee \beta)}_{*}$	1, 3 IA
1	(5)	$\overbrace{\neg\alpha}^{*}$	2, 4 RAA
6	(6)	β	Assumption
6	(7)	$\alpha \vee \beta$	6 I$\vee$
1, 6	(8)	$\underbrace{(\alpha \vee \beta) \wedge \neg(\alpha \vee \beta)}_{*}$	1, 7 IA
1	(9)	$\overbrace{\neg\beta}^{*}$	6, 8 RAA
1	(10)	$\underbrace{\neg\alpha \wedge \neg\beta}_{*}$	5, 9 I$\wedge$

$$\frac{\Gamma \vdash \overbrace{\neg\alpha \wedge \neg\beta}^{(*)}}{\Gamma \vdash \underbrace{\neg(\alpha \vee \beta)}_{(*)}}$$

Proof (non-starred version)

1	(1)	$\neg\alpha \wedge \neg\beta$	Assumption
2	(2)	$\underbrace{\alpha \vee \beta}_{*}$	Assumption
3	(3)	α^{*}	Assumption
4	(4)	β^{*}	Assumption
1	(5)	$\neg\alpha$	1 E$\wedge$
1, 3	(6)	$\underbrace{\alpha \wedge \neg\alpha}_{*}$	3, 5 IA
3	(7)	$\underbrace{\neg(\neg\alpha \wedge \neg\beta)}_{*}$	1, 6 RAA
1	(8)	$\neg\beta$	1 E$\wedge$
1, 4	(9)	$\underbrace{\beta \wedge \neg\beta}_{*}$	4, 8 IA
4	(10)	$\underbrace{\neg(\neg\alpha \wedge \neg\beta)}_{*}$	1, 9 RAA
2	(11)	$\underbrace{\neg(\neg\alpha \wedge \neg\beta)}_{*}$	2, 3, 4, 7, 10 E$\vee$
1, 2	(12)	$\underbrace{(\neg\alpha \wedge \neg\beta) \wedge \neg(\neg\alpha \wedge \neg\beta)}_{*}$	1, 11 IA
1	(13)	$\neg(\alpha \vee \beta)$	2, 12 RAA

Proof (non-starred version)

1	(1)	$\overbrace{\neg\alpha \wedge \neg\beta}^{*}$	Assumption
2	(2)	$\alpha \vee \beta$	Assumption
3	(3)	α	Assumption
4	(4)	β	Assumption
1	(5)	$\overbrace{\neg\alpha}^{*}$	1 E$\wedge$
1, 3	(6)	$\underbrace{\alpha \wedge \neg\alpha}_{*}$	3, 5 IA
3	(7)	$\neg(\neg\alpha \wedge \neg\beta)$	1, 6 RAA
1	(8)	$\overbrace{\neg\beta}^{*}$	1 E$\wedge$
1, 4	(9)	$\underbrace{\beta \wedge \neg\beta}_{*}$	4, 8 IA
4	(10)	$\neg(\neg\alpha \wedge \neg\beta)$	1, 9 RAA
2	(11)	$\neg(\neg\alpha \wedge \neg\beta)$	2, 3, 4, 7, 10 E$\vee$
1, 2	(12)	$\underbrace{(\neg\alpha \wedge \neg\beta) \wedge \neg(\neg\alpha \wedge \neg\beta)}_{*}$	1, 11 IA
1	(13)	$\underbrace{\neg(\alpha \vee \beta)}_{*}$	2, 12 RAA

Law of Non-Contradiction (LNC)

$$\Gamma \vdash \neg(\alpha \wedge \neg\alpha)$$

Proof

1	(1)	$\overbrace{\alpha \wedge \neg\alpha}^{*}$	Assumption
1	(2)	$\underbrace{\alpha \wedge \neg\alpha}_{*}$	1 R
	(3)	$\neg(\alpha \wedge \neg\alpha)$	1, 2 RAA

Modus Ponens (MPP)

$$\frac{\Gamma \vdash \alpha^* \qquad \Delta \vdash \neg\alpha \vee \beta}{\Gamma, \Delta \vdash \beta}$$

Proof

1	(1)	α^*	Assumption
2	(2)	$\neg\alpha \vee \beta$	Assumption
3	(3)	$\underbrace{\neg\beta}_{*}$	Assumption
1, 3	(4)	$\underbrace{\alpha \wedge \neg\beta}_{*}$	1, 3 I∧
1, 3	(5)	$\underbrace{\neg(\neg\alpha \vee \beta)}_{*}$	4 De Morgan
1, 2, 3	(6)	$\underbrace{(\neg\alpha \vee \beta) \wedge \neg(\neg\alpha \vee \beta)}_{*}$	2, 5 IA
1, 2	(7)	β	3, 6 RAA

$$\frac{\Gamma \vdash \alpha \qquad \Delta \vdash \overbrace{\neg\alpha \vee \beta}^{*}}{\Gamma, \Delta \vdash \beta^*}$$

Proof

1	(1)	α	Assumption
2	(2)	$\underbrace{\neg\alpha \vee \beta}_{*}$	Assumption
3	(3)	$\neg\beta$	Assumption
1, 3	(4)	$\alpha \wedge \neg\beta$	1, 3 I∧
1, 3	(5)	$\neg(\neg\alpha \vee \beta)$	4 De Morgan
1, 2, 3	(6)	$\underbrace{(\neg\alpha \vee \beta) \wedge \neg(\neg\alpha \vee \beta)}_{*}$	2, 5 IA
1, 2	(7)	β^*	3, 6 RAA

Elimination of Absurd (EA)

$$\frac{\Gamma \vdash \overbrace{\alpha \wedge \neg\alpha}^{*}}{\Gamma \vdash \beta^{*}}$$

Proof

1	(1)	$\overbrace{\alpha \wedge \neg\alpha}^{*}$	Assumption
1	(2)	α^{*}	1 E$\wedge$
1	(3)	α	2 R
1	(4)	$\underbrace{\neg\alpha}_{*}$	1 E$\wedge$
1	(5)	$\underbrace{\neg\alpha \vee \beta}_{*}$	4 I$\vee$
1	(6)	β^{*}	3, 5 MPP

Material conditional

$$\frac{\Gamma, \alpha^{(*)} \vdash \beta^{(*)}}{\Gamma \vdash \neg\alpha \vee \beta}$$

Proof (non-starred version)

1	(1)	$\alpha \vdash \beta$	Premise
	(2)	$\alpha \vee \neg\alpha$	LEM
3	(3)	α	Assumption
1, 3	(4)	β	3 Premise
1, 3	(5)	$\neg\alpha \vee \beta$	4 I$\vee$
6	(6)	$\neg\alpha$	Assumption
6	(7)	$\neg\alpha \vee \beta$	6 I$\vee$
1	(8)	$\neg\alpha \vee \beta$	2, 3, 5, 6, 7 E$\vee$

Proof (starred version)

1	(1)	$\alpha^* \vdash \beta^*$	Premise
	(2)	$\alpha \vee \neg\alpha$	LEM
3	(3)	α^*	Assumption
3	(4)	α	3 R
1, 3	(5)	β^*	3 Premise
1, 3	(6)	β	5 R
1, 3	(7)	$\neg\alpha \vee \beta$	6 I$\vee$
8	(8)	$\neg\alpha$	Assumption
8	(9)	$\neg\alpha \vee \beta$	8 I$\vee$
1	(10)	$\neg\alpha \vee \beta$	2, 4, 7, 8, 9 E$\vee$

$$\frac{\Gamma, \alpha^* \vdash \beta}{\Gamma \vdash \neg\alpha \vee \beta}$$

Proof

1	(1)	$\alpha^* \vdash \beta$	Premise
	(2)	$\alpha \vee \neg\alpha$	LEM
3	(3)	α^*	Assumption
3	(4)	α	3 R
1, 3	(5)	β	3 Premise
1, 3	(6)	$\neg\alpha \vee \beta$	5 I$\vee$
7	(7)	$\neg\alpha$	Assumption
7	(8)	$\neg\alpha \vee \beta$	7 I$\vee$
1	(9)	$\neg\alpha \vee \beta$	2, 4, 6, 7, 8 E$\vee$

$$\frac{\Gamma, \alpha \vdash \beta^*}{\Gamma \vdash \underbrace{\neg\alpha \vee \beta}_{*}}$$

Proof

1	(1)	$\alpha \vdash \beta^*$	Premise
	(2)	$\alpha \vee \neg\alpha$	LEM
3	(3)	α	Assumption
1, 3	(4)	β^*	3 Premise
1, 3	(5)	β	4 R
1, 3	(6)	$\neg\alpha \vee \beta$	5 I$\vee$
7	(7)	$\neg\alpha$	Assumption
7	(8)	$\neg\alpha \vee \beta$	7 I$\vee$
1	(9)	$\neg\alpha \vee \beta$	2, 3, 6, 7, 8 E$\vee$

T-scheme derived rules

$$\frac{\Gamma \vdash \overbrace{\neg T\underline{\alpha}}^{(*)}}{\Gamma \vdash \underbrace{T\underline{\neg\alpha}}_{(*)}}$$

Proof (non-starred version)

1	(1)	$\neg T\underline{\alpha}$	Assumption
2	(2)	$\underbrace{\neg T\underline{\neg\alpha}}_{*}$	Assumption
1	(3)	$\neg\alpha$	1 T-scheme
2	(4)	$\underbrace{\neg\neg\alpha}_{*}$	2 T-scheme
2	(5)	α^*	4 DN
1, 2	(6)	$\underbrace{\alpha \wedge \neg\alpha}_{*}$	3, 5 IA
1	(7)	$\neg\neg T\underline{\neg\alpha}$	2, 6 RAA
1	(8)	$T\underline{\neg\alpha}$	7 DN

Proof (starred version)

1	(1)	$\overbrace{\neg T\underline{\alpha}}^{*}$	Assumption
2	(2)	$\neg T\underline{\neg\alpha}$	Assumption
1	(3)	$\underbrace{\neg\alpha}_{*}$	1 T-scheme
2	(4)	$\neg\neg\alpha$	2 T-scheme
1, 2	(5)	$\overbrace{\neg\alpha \wedge \neg\neg\alpha}^{*}$	3, 4 IA
1	(6)	$\underbrace{\neg\neg T\underline{\neg\alpha}}_{*}$	2, 5 RAA
1	(7)	$\underbrace{T\underline{\neg\alpha}}_{*}$	7 DN

$$\frac{\Gamma \vdash \overbrace{T\underline{\neg\alpha}}^{(*)}}{\Gamma \vdash \underbrace{\neg T\underline{\alpha}}_{(*)}}$$

Proof (non-starred version)

1	(1)	$T\underline{\neg\alpha}$	Assumption
2	(2)	$\overbrace{T\underline{\alpha}}^{*}$	Assumption
2	(3)	α^*	2 T-scheme
1	(4)	$\neg\alpha$	1 T-scheme
1, 2	(5)	$\overbrace{\alpha \wedge \neg\alpha}^{*}$	3, 4 IA
1	(6)	$\neg T\underline{\alpha}$	2, 5 RAA

Proof (starred version)

1	(1)	$\overbrace{T\underline{\neg\alpha}}^{*}$	Assumption
2	(2)	$T\underline{\alpha}$	Assumption
2	(3)	α	2 T-scheme
1	(4)	$\overbrace{\neg\alpha}^{*}$	1 T-scheme
1, 2	(5)	$\overbrace{\alpha \wedge \neg\alpha}^{*}$	3, 4 IA
1	(6)	$\underbrace{\neg T\underline{\alpha}}_{*}$	2, 5 RAA

2.5.2 Proofs of the liar paradox

Let λ be the liar sentence:

$$(\lambda) \quad \neg T\underline{\lambda}$$

Then, in DLEAC λ is a dialetheia.

Proof

1	(1)	$T\underline{\lambda}$	Assumption
1	(2)	$\neg T\underline{\lambda}$	T-scheme
	(3)	$\neg T\underline{\lambda} \vee \neg T\underline{\lambda}$	Material conditional
	(4)	$\neg T\underline{\lambda}$	Assumption
	(5)	$\neg T\underline{\lambda}$	4 R
	(6)	$\neg T\underline{\lambda}$	3, 4, 5, 4, 5 E$\vee$
	(7)	$T\underline{\lambda}$	T-scheme
	(8)	$\neg T\underline{\lambda} \wedge T\underline{\lambda}$	I$\wedge$

Chapter 3

Gluon theory and its consequences

It all starts with the following self-evident truth: rational agents are able to individually refer to objects, be they simple (i.e. atomic) or composite (i.e. made up of parts). It is quite trivial to note we have the linguistic resources to point out specific objects. For instance, it is one object that I refer to by means of the following expressions: «the Brooklyn Bridge» (a proper name), «the king of France who ruled from May 14, 1643, to September 1, 1715» (a definite description), «my computer» (a combination of an indexical and a predicate), «a quark» (an indefinite description), etc. And this cognitive ability is indeed fundamental, since it appears to be embroiled in many philosophically relevant concepts, such as being, intentionality, identity, and others.

The problem arises when we note that we can perceive something as an object – i.e. one thing – even when it is made up of parts. For that very object is one, since it is an object and you can individually refer to it; but it is also many, since it is composed by a collection of other objects – its parts. Hence the urge to answer these questions: how do the parts (metaphysically) form a whole? What (metaphysically) accounts for the unity of the whole? This, essentially, is the so called *one and the many problem*.

Solving this problem does not seem to be an easy task, though. Various solutions have been proposed but, according to Priest, all these solutions run into the vicious Bradley regress. To make clear what this regress is, let us call that which constitutes the parts as a single unit 'gluon'. Suppose we have a unity made up of p_1 and p_2. Then, the gluon binds them together. But now, what binds p_1, p_2 and the gluon together? This must be some other gluon, or super-gluon. And now we are off to a vicious infinite regress: «[o]ur original problem was how a unity of parts is possible. We need an explanation. Given a bunch of parts, simply invoking another object does not do this» (Priest, 2014b, p. 11).

In Priest's presentation of the issue it is possible to see a glimpse of his proposed solution, which goes in the direction of unconventional types of objects:

> Here, then, is our problem of unity. [...] Take any thing, object, entity, with parts, $p_1, \ldots, p_n$. [...] A thing is not merely a plurality of parts: it is a unity. There must, therefore, be something which constitutes them as a single thing, a unity. Let us call it, neutrally [...] the gluon of the object, g. Now what of this gluon? Ask whether it itself is a thing, object, entity? It both is and is not. It is, since we have just talked about it, referred to it, thought about it. But it is not, since, if

> it is, $p_1, \ldots, p_n, g$, would appear to form a congeries, a plurality, just as much as the original one. If its behaviour is to provide an explanation of unity, it cannot simply be an object.
>
> Priest (2014b, p. 9)

Thus, it is because the gluon is an object – i.e. it is because of its objecthood – that it runs into the Bradley regress and does not solve the problem.

Priest's attempt to find a solution gives rise to a very interesting as well as unorthodox theory called gluon theory, whose main reference is Priest (2014b). Such a theory can be thought as composed of some major ingredients, namely: dialetheism, noneism, mereology and a very peculiar identity relation. Put them together, add some appropriate definition and a specific account of properties, and that's it.

In this chapter I will first present gluon theory from the bottom, introducing all its major (and others, e.g. plural quantification) ingredients – except dialetheism, for which I refer to §1. Then, I will examine what I believe to be some undesired consequences of Priest's theory which – I think – significantly decrease its rational acceptability. For that, I will follow and extend Carrara, Mancini, and Smid (2021).

3.1 Noneism

Noneism[1] – also known as modal meinongianism – is the view that there are nonexistent objects. It was developed by Priest (2016) and Berto (2008) starting from Richard Routley's work, as an improvement of Alexius Meinong's theory of objects that deploys modality (with both possible and impossible worlds) to overcome the obstacles caused by his unconstrained characterisation principle (CP).

There are some reasons to allow for nonexistent objects in the philosophical theorizing.[2] Among them, the problem of intentionality may count as one of the most important motivations for thinking there are nonexistent objects. For according to a well-known principle called "principle of intentionality", many mental *phenomena* are characterized by an intentional directedness towards an object. For instance, to desire is always to desire something, to fear is always to fear something, and so forth. The problem is that sometimes people desire, fear and experience other mental acts upon things that do not exist, and this seems to put such a principle in a difficult position. Nevertheless, many philosophers found it too appealing to be given up

[1] It was Richard Routley in his *Exploring Meinong's Jungle* (Routley, 1980) who coined this name. To be precise, noneism is Routley's (simpler) version of the meinongian theory of objects, that was later refined by Priest and Berto by implementing a particular kind of modality – hence the name *modal meinongianism*.

[2] For example, nonexistent objects seem to have a part to play in some kinds of discourse, such as fictional discourse and that about the past and the future. Moreover, if we banish nonexistent objects it does not seem we can easily make sense of the truth of some negative singular existence sentences such as «Pegasus does not exist». In short: the claim that there are nonexistent objects does appear to have a considerable explanatory force.

completely and followed the strategy initiated by Meinong: there is indeed an object for every mental state whatsoever, if not an existent object then at least a nonexistent one.

However, nonexistent objects are suspicious and generally rejected, and this is mainly because of the following problem:

> to be able to say truly of an object that it doesn't exist, it seems that one has to presuppose that it exists, for doesn't a thing have to exist if we are to make a true claim about it?
>
> Reicher (2019)

The generally accepted answer to this question is «Yes». For according to a very well-established tradition which dates back to Frege (1952a) and Quine (2011) – among others –, existence is not a predicate, but the result of a quantification occurring in a meaningful sentence.[3] Thus, we just can't speak truly about nonexistent objects: if we could, we would make them existent.[4] Therefore, they fall outside the domain of quantification and no variable can range over them. But this view does not seem to be able to account for the legitimacy of many meaningful and possibly true statements involving nonexistent objects. For example, consider:

(a) Sancho Panza is Don Quixote's faithful squire.

(b) Voldemort does not exist.

(c) On June 22nd 2021, the one who wrote this thesis imagined a golden mountain.

(a) and (b) are supposedly true, while (c) is definitely true (believe me!). But they all contain terms denoting nonexistent objects: some fictional or imaginary objects. Thus, (a), (b) and (c) may count as counterexamples to the quinean approach.[5]

That said, there are several ways to try addressing this problem. For example, to resolve (a) we can just go against our intuition and say that it is false instead of true; or we can rely on a "story operator" to make (a) true according to the story the novel tells.[6] Again, for (a) and (b) we can assume Russell's theory of definite

[3]To this regard, recall the quinean famous slogan «To be is to be the value of a variable» (*Ibid.*) that needs no introduction.

[4]Note that a dialetheist may find appealing to accept that nonexistent objects are also existent objects, i.e. that are dialetheic (or contradictory, or inconsistent) objects. For now, let's remain consistent and keep noneism and dialetheism separate.

[5]These examples do not exhaust the list of situations where a true sentence with a quantification over nonexistent objects seems to be involved. To mention just a few other cases, consider the analytic truth «The round square is round and square» where «round square» denotes a nonexistent object, and the sentences containing terms which refer to objects of the past (or future).

[6]See for example Künne (2011).

descriptions and try to paraphrase away every proper names.[7] (In this perspective, a proper name is just a linguistic shortcut of a definite description.) Thus, if we implement this strategy we can show that these sentences only resort to terms which refer to (existent) objects and that are unproblematic, as well as true. Otherwise, another different approach is to accept a positive free logic, where singular terms do not need to denote anything in order to be meaningful and the sentences containing non-denoting singular terms can be true.[8] Unfortunately, these strategies do not provide a generally accepted solution to the problem, which is still open. But now, a further and very straightforward approach is to allow for nonexistent objects. This require (i) to include such a kind of objects in the domain, (ii) to take existence to be a predicate (usually represented by E), and (iii) to consider quantification to be neutral – that is, not ontologically committing (*contra* Quine). Moreover, this view comes in some different versions. In this regard, the major reference is Meinong's theory of objects (Meinong, 1960), but there are also many contemporary theories[9] that attempt to make sense of (alleged) talk about nonexistent objects and (seeming) intentional directedness to nonexistent objects (as in (c), the above example). Noneism is among these.

To present noneism, let's start with Meinong and the object theory he developed to account for (the semantics of) intentionality, i.e. the feature of cognition whereby a mental state is "directed towards" an object of some kind. He calls «object» whatever can be experienced in some way – i.e. be the target of a mental act. Thus, everything is an object: the category of objects has no complement – this is sometimes called *umbrella view*.[10] Furthermore, he distinguishes two feature of an object: its being (*Sein*) and its so-being (*Sosein*). The being of an object is its existential status. An arbitrary object can have being or can have no being (these two categories are meant to be exhaustive and exclusive). If it has being, it can exist (i.e. it is a real object, like concrete objects and those linked with time in general) or subsist (i.e. it is an ideal object, like timeless abstract objects), being existence and subsistence different modes of being.[11] Instead, objects with no being, such as the golden mountain, the greatest prime number and the round square, neither exist nor subsist.

Now, according to Meinong objects abide by the following principle:

- the (unrestricted) *characterization principle*, (CP): if an object is characterized in a certain way, it has its characterizing properties.

[7] This strategy was suggested, for example, by Quine (1948).

[8] See Crane (2012) for an in-depth analysis of this approach.

[9] For example, those which implement a de-ontologization strategy (e.g. Crane, 2012, 2013), those which opt for fictionalism (see Eklund, 2019 for an introduction) and the neo-Meinongian theories (e.g. Jacquette, 2011, Routley, 1980 and Zalta, 2012).

[10] Actually, this claim can be challenged on the basis of some Meinong's lecture notes from 1917/18. Here, he discusses some cases of so-called *defective* objects, i.e. objects that involve paradoxes like the liar, which may fall outside the category of objects. As we will see in the next sections, Priest's dialetheism can help shedding light on this topic – i.e. non-objecthood.

[11] To be precise, the relation between Meinong's existence and subsistence is more complex: real objects both exist and subsist, whereas ideal objects only subsist.

Thus, given any property or collection of properties represented by the open formula $A(x)$, there is (to be read without any existential commitment) an object a such that $A(a)$ is true. However, it is easy to see that this unrestricted version of CP is deeply problematic. For take $A(x)$ to be $B(x) \wedge Ex$, where $B(x)$ is any open formula (i.e. it represents any property whatsoever) and E is the predicate of existence. Then, if we apply CP we get that whatever object actually exists, e.g. even abstract and contradictory objects. Even worse than that, take $A(x)$ to be $B(x) \wedge C$, where C is any closed formula. By applying CP we immediately get triviality.

To avoid these problems – which are usually gathered together under the name *characterization problem* – Meinong assumes another principle:

- the *indifference principle*: being and non-being are not part of an object's nature, i.e. its characterizing properties, but nevertheless «the law of excluded middle lays it down that every object necessarily stands in a fact of being or in a fact of non-being» (Findlay, 1963).

Thus, an object's so-being is independent of its being: the existential status of an object and what properties it has are quite separate issues. That is, E cannot occur in $A(x)$ since it is not a characterizing (or *nuclear*) property, so that the problem we have mentioned before vanishes. Nevertheless, many issues remain unsolved. In principle, we are allowed to restrict CP to get rid of suspicious and unacceptable objects; but then, in order to escape any charge of making an *ad hoc* move we are required to ground such a restriction independently from this benefit. And this has not been yet accomplished. As Priest puts it:

> no noneist has even accepted the CP in its pristine [unrestricted] form. The standard response, from Meinong onwards, has been to accept it only if the properties deployed in the CP are of a certain kind: *assumptible*, *characterizing*, *nuclear*, the names vary. [...] The problem for this line is to give a principled of what constitutes a characterizing predicate and why. No one, as far as I am aware, has been able to do this.
>
> Priest (2016, p. 83)

Moreover, Priest argues that any kind of restriction to CP will always be completely unjustified:

> [I]t would appear to be the case that we can think about an object satisfying any set of conditions whatsoever. Phenomenologically, at least, there is absolutely no

> difference between contemplating an object that has only officially characterizing properties—whatever those are—and one that has some non-characterizing properties as well, say existence. [...] *Qua* object of thought, each object seems to have all the properties deployed. Drawing distinctions within these properties seems entirely unmotivated.
>
> Priest (2016, pp. 83-84)

Then, here is the solution he proposes:

> I suggest, the object characterized by a representation has the characterizing properties, not necessarily in the actual world, but in the worlds (partially) described by the relevant representation. [...] In this way, the CP can be accepted *in full generality*: we just do not assume that an object characterized in a certain way has its characterizing properties at the actual world, only at the worlds which realize the way the agent represents things to be in the case at hand.
>
> Priest (2016, p. 84)

Thus, Priest locates the issue of nonexistent objects in its natural Meinongian home, intentionality, and uses the 'technology' of possible and impossible worlds to frame an appropriate account. Let's now briefly explore the basic technical details of his framework.[12]

Take first-order constant domain[13] $S5$ modal logic as the starting point for the semantics of intentional operators. Then, augment its language with a set of intentional operators, Ψ_i (e.g. Ψ_1 might be «knows that», Ψ_2 might be «fears that», etc.), such that if t is any term and α is any formula, $t\Psi_i\alpha$ is a formula. Now, the semantics of Ψ_i is a simple generalization of the binary-relation semantics for modal operators. Thus, for any intentional verb, Ψ_i, its denotation is a function that maps each member of the domain, D, to a binary relation on the set of worlds, W. Let's call $R^{d(t)}_{\Psi_i}$ such a relation – i.e. the accessibility relation. Of course, $R^{d(t)}_{\Psi_i}$ depends on the intensional operator we are considering, Ψ_i, as well as on the individual (i.e. the agent) who performs the corresponding mental act, $d(t)$ – i.e. the denotation of the

[12]In what follows, I shall limit myself to only the major aspects of Priest's account for intentional operators. For instance, I will not explain how he regiments the identity relation to make it violates the principle SI – i.e. substitutivity of identicals – in intentional contexts. I will not even describe how his framework accounts for what he calls intentional predicates, i.e. intentional verbs with noun-phrase complements, as in «Homer worshipped Zeus». And I'll skip the semantics for definite and indefinite descriptions. The reader who is interested in having more details on that can refer to Priest (2016, §§1-4).

[13]According to Priest, the reason why constant domain worlds would be better than variable domain worlds corresponds to the very same endorsement of noneism. Quoting his words: «[t]he major reason why, it is usually assumed, variable domains are more appropriate for a world semantics than constant domain, is that it seems clear that different things exist at different worlds. [...] But this is to assume that the denizens of a world's domain are precisely the things that exist there. And this is rejected by noneism. If one is a noneist, there would seem to be no reason why the domain of each world should not be exactly the same, namely the set of all objects—whatever an object's existential status at that world» (Priest, 2016, p. 13).

term t. Therefore, for any t and α the truth and falsity conditions for $t\Psi_i\alpha$ in a world w are:

$$(\Psi_i 1) \qquad 1 \in \nu_w(t\Psi_i\alpha) \text{ iff for all } w' \text{ such that } wR^{d(t)}_{\Psi_i}w',\ 1 \in \nu_{w'}(\alpha)$$

$$(\Psi_i 2) \qquad 0 \in \nu_w(t\Psi_i\alpha) \text{ iff for some } w' \text{ such that } wR^{d(t)}_{\Psi_i}w',\ 0 \in \nu_{w'}(\alpha)$$

Then, to properly model a particular intensional operator we just need to impose the right constraints on the corresponding accessibility relation (e.g. that it is reflexive, etc).

We now have to make two further changes to this quite standard logical setting. The first one is precisely to turn to noneism by including both existent and non-existent objects in the domain. To do that, we need to remove any (alleged) ontological commitment from quantification. Then, we can quantify over objects by using the particular and the universal quantifiers, $\mathfrak{S}$ and $\mathfrak{A}$ respectively. $\mathfrak{S}xPx$ has to be read as «Some x is such that Px», in contrast to the standard reading of $\exists xPx$ as «There exists/there is an x such that Px».[14] In addition to this, we have to establish an existence predicate, E, to signify that the object exists. Thus, if we want to say that there exists something that is P, we should write $\mathfrak{S}x(Ex \wedge Px)$.[15]

The second change we have to make is to expand the worlds apparatus by adding so-called non-normal or impossible worlds.[16] These worlds are meant to be logically impossible: their logics differ somewhat from the logic of the actual (and other possible) world(s). (Of course, what the logic of the actual world is becomes a central issue. But there is no need to face it here and we can just move forward.) Let's call I the set of impossible worlds and P that of possible worlds. We then impose the following constraints: $P \cup I = W$ (exhaustivity), $P \cap I = \varnothing$ (exclusivity) and $@ \in P$. For any $w \in I$, all complex formulas are treated as if they were atoms, in that they are related to truth values directly, not recursively. Thus, we expand the notion of interpretation by including both P and I – instead of just W – in the corresponding tuple, and by

[14] Priest takes «there exists» and «there is» as synonyms. Even if I won't go into this matter here, it should be said that this point can be questioned. For more on that see e.g. Priest (2016, §17).

[15] For the record, Priest considers the notion of existence so fundamental to our thought that unlikely we can provide for a definition. It seems to resist explanation in any but circular terms and we should consider it a primitive notion. Nevertheless, he admits a guiding *criterion* – though not a definition – for predicating the existence of an object: «to exist is to have the potential to interact causally» (Priest, 2016, p. xxviii). Moreover, likewise Routley (1980) he takes existence the only mode of being. That is, there are not subsistent objects: lacking any spatio-temporal characterization – and consequently any causal power – abstract objects are nonexistent in the same way as fictional and impossible objects are.

[16] To this regard, an important remark is to be made. Beyond possible and impossible worlds, Priest (2016)'s semantics for intentional operators also distinguishes between open and closed worlds. I don't enter the details of such a distinction here. Suffice it to say that Priest's notion of impossible world is narrower than the one I refer to. For he takes impossible worlds as worlds where only conditionals may behave arbitrarily, and open worlds as worlds where all formulas may behave arbitrarily – they can be "completely anarchic". Thus, Priest's impossible worlds are just special cases of open worlds. Instead, when I say «impossible worlds» I mean Priest's open worlds. Given the introductory purpose of this section, I've decided to leave this further classification behind and resort only to the most comprehensive class of worlds – in my jargon, that of impossible worlds – for the sake of simplicity.

tuning the semantic clauses for connectives and operators accordingly.[17] Finally, validity is truth preservation at @ in all interpretations and logical consequence is defined in the usual way.

Now, there are several reasons why we might want to go this way. For instance, Priest shows that with this possible and impossible worlds semantics we can get a relevant logic – and therefore a strict conditional – without resorting to the Routley-Meyer ternary relation, so that we can overcome the well-known fallacies of relevance with a more standard binary-relation semantics. Furthermore, impossible worlds enable us to account for hyper-intensional contexts,[18] such as those related to some propositional attitudes (e.g. imagination – see Berto, 2017), and solve some issues with epistemic and doxastic operators which have been often grouped under the label of "logical omniscience". But the most significant motivation here is that the inclusion of impossible worlds provides an elegant solution to the characterization problem.

Priest's move with respect to CP is to accept it in full generality, that is with no restriction on the class of the characterizing properties. What we need to add in order to avoid the problems mentioned above (that is, to solve the characterization problem) is the following idea:

> [...] we just do not assume that an object characterized in a certain way has its characterizing properties at the actual world, only at the worlds which realize the way the agent represents things to be in the case at hand.
>
> Priest (2016, p. 85)

The insight here is that whenever a cognitive agent characterizes an object as having certain properties, she is performing an intentional act: that of representing an object *via* some properties. Thus, we can apply the semantics we have just described and prevent both the 'actualization' of every object and triviality thanks to modality – i.e. the worlds framework. To see that, let Φ be an intentional operator of the form «... represents ... as holding [such and such properties]». Now:

> let $A(x)$ be any condition; someone can intend an object of thought characterized by $A(x)$, and let 'c_A' rigidly designate it. [...] Then we may not have $@ \Vdash^+ A(c_A)$, but if a is the relevant agent, and Φ is the appropriate intentional operator, we do have $@ \Vdash^+ a\Phi A(c_A)$; so at every w such that $@R_\Phi^{d(a)}w$, $w \Vdash^+ A(c_A)$.
>
> Priest (2016, p. 85)

[17]The only operators whose semantic clauses would refer to impossible worlds are the intensional ones.

[18]Hyper-intensional contexts are those contexts where the principle of substitution *salva veritate* of expressions with the same intension fails.

As a result, CP seems to be completely liberalized: we do not need to put any limitation on it.

Finally, two additional clarifications. The first one is that, since we can characterize an object inconsistently, we must have inconsistent worlds – which we have, i.e. impossible worlds. For example, the round square and the four sides triangle do not exist in any $w \in P$, but they are meant to be legitimate denizens of some $w \in I$. The second one is that an object characterized in a certain way has not only its characterizing properties at the appropriate worlds, but also those that follow from them. For example, if Voldemort has exactly two eyes in the worlds which realize J.K. Rowling's novels, then Voldemort has an even number of eyes there (provided number theory is canonical in such worlds).

Now, Priest's framework does seem to fulfil the purpose it was conceived for. Though, it appears to have some deficiencies. For instance, Priest remains silent about which properties are existence-entailing and which are not; and it is not clear which properties nonexistent objects have in the actual world. Several other objections have been made to noneism and the semantics for intentional operators we have just presented. To one of these I turn in the next section, before we move to the next 'ingredient' of gluon theory – i.e. mereology.

3.1.1 Coda on noneism: objecting to an objection Priest already objected

I have to admit that Priest's noneism strikes me as quite appealing (perhaps *modulo* the multiple denotation semantics he develops to solve the thorny paradox of denotation).[19] It is a simple and intuitive view and, for what I can see, its competitors (e.g. Russell and Quine's perspectives) cannot do better with respect to the problems posed by intentionality and the topics related to nonexistent objects. Moreover, it is a very general theory and can be coherent with a broad spectrum of more specific views (e.g. platonism). However, there is a possible weakness that struck me the first time I thought about the way Priest uses modality to liberalize CP. After I realized there might have been such problem, I found the very same objection in Beall (2006), and later I came across Priest's reply in Priest (2016, §13.6.3). This point, I think, is worth discussing here, despite being a bit off-topic with respect to the main purpose

[19] There is a paradox in the family of the semantic paradoxes of self-reference known as the paradox of Hilbert and Bernays. It stems from the naive notion of denotation function, d, which is very intuitive but provably inconsistency-producing. d is governed by the principle we can call the d-schema: $d(\underline{t}, x) \dashv\vdash t = x$, where t is any term. Now, likewise Tarski and his theorem concerning the truth predicate, Hilbert and Bernays used their paradox to draw the conclusion that a consistent theory cannot contain its own denotation function. Instead, Priest opts for a different strategy. He does not impose any restriction on the denotation function and maintains the d-schema in full generality. But in order to avoid the paradox, he assumes that any term may have more than one denotation. This move corresponds to the very same strategy Priest applied to the liar paradox and its solution, i.e. dialetheism. In the same way that dialetheism solve the liar paradox by allowing sentences to have possibly more than one truth-value, the multiple denotation semantics solve the paradox of Hilbert and Bernays by allowing a term to have possibly more than one denotation. The most significant reason why I am not very sympathetic towards this solution is that it requires a non-transitive identity relation. More on that in what follows. See Priest (2016, §8) for the details.

of this work. My aim is to give a contribute – even if only in terms of clarification – to this matter, and the present section will prove a defense of Priest's semantics.

As we explained, noneism liberalizes CP and deploys a possible and impossible worlds semantics to avoid the unpleasant drawbacks implied by such an unrestricted version of this principle. By doing so, Priest says we don't need to distinguish between characterizing (or nuclear, ...) and non-characterizing (or non-nuclear, ...) properties, since any object has whatever property it is characterized as having, just not necessarily in the actual world. Thus, take S for «being square», R for «being round» and E for «being» – i.e. «existent». Then, $\mathfrak{S}x(Rx \wedge Sx \wedge Ex)$. But such an existent round square does not exist in the actual world; it exists in the logically impossible worlds which correspond to its representation. Now, Beall (2006) argues that this solution does not always work; for example, when we include the actuality operator, $\mathcal{A}$, into the language. Let me introduce the (alleged) problem by quoting Priest's words:

> *Objection*: Let $\mathcal{A}$ be the operator *it is actually the case that*. [...] Now, consider the condition $x = x \wedge \mathcal{A}B$, for an arbitrary B. Let b be the object characterized by this condition. Then at some world w, $b = b \wedge \mathcal{A}B$. Hence, B is actually the case. [Hence, actual triviality.]
>
> Priest (2016, pp. 247-248)

Priest's reply is based on the following insight: «[t]he fact that $\mathcal{A}B$ holds at w, does not imply that $\mathcal{A}B$, and so B, is true at @» (Priest, 2016, p. 248). For there are genuine counterexamples to Beall's argument: at some impossible worlds, $\mathcal{A}B$ may hold without B holding at @. That is to say that the following truth condition of $\mathcal{A}$ are incorrect: for any $w \in W$, $w \Vdash^+ \mathcal{A}B$ iff $@ \Vdash^+ B$. In other words, the problem with Beall's line of reasoning – and, I must confess, also with the way I initially looked at this issue – is that, given two arbitrary worlds, $w_1 \in W$ and $w_2 \in W$, such that $w_1 \neq w_2$, you are not allowed to draw any conclusion at w_2 from some true formula at w_1 – even if that formula predicates something about the world w_1. And this is so precisely because impossible worlds inhabit Priest's semantics, so that there is no common logic shared by all the worlds in W – which is exactly what you need to make Beall's objection convincing.

Though, I had some reservations about Priest's reply at first. Let me try to spell out what I had in mind. Call **Act** the truth condition of $\mathcal{A}$ we have just mentioned: for any $w \in W$, $w \Vdash^+ \mathcal{A}B$ iff $@ \Vdash^+ B$. I agree **Act** is not correct for any $w \in W$, but it is correct for a specific subset of worlds: P. That is, **Act** works properly for any possible world. And Priest thinks the same: «[t]he truth conditions given by **Act** are correct if w is a possible world» (Priest, 2016, p. 249). The reason is simple: by definition, possible worlds do share the same logic. Now, let us call **Act*** the restriction of **Act** to possible worlds: for any $w \in P$, $w \Vdash^+ \mathcal{A}B$ iff $@ \Vdash^+ B$. Consider again the condition $x = x \wedge \mathcal{A}B$, for an arbitrary B, and let b be the object characterized in this way. Given

Act*, we just need (at least) one possible world where $b = b \wedge \mathcal{A}B$ holds to get in trouble. And I couldn't see any reason why this should have not been the case. For the condition $x = x$ is logically valid for any $w \in P$. Therefore, there should always be (at least) one possible world such that $b = b \wedge \mathcal{A}B$.

However, this argument is mistaken. For is it possible for $\mathcal{A}B$ to hold in a possible world, whatever B is? No, and the reason is **Act***. Let's make an example. Let B represents «Filippo Mancini is a woman» and $C(x)$ the condition $x = x \wedge \mathcal{A}B$. If we accept **Act***,[20] since B is (only) false at @, the same is for $\mathcal{A}B$ at any possible worlds. In other words, $b = b \wedge \mathcal{A}B$ can hold only at impossible worlds: due to **Act***, there is no possible world such that $b = b \wedge \mathcal{A}B$. Thus, Priest is right and the problem vanishes.[21]

3.2 Mereology

Mereology is the study of parthood relation, <.[22] Many philosophers consider such a relation to be fundamental, meaning that it is one of the basic irreducible relations among the entities that constitute the ontology of our world. Then, they ascribe it a very significant metaphysical import that makes it deserve a deep investigation.

The origin of mereology can be traced back to both Husserl (1901)'s *Logical Investigation* and Lesniewski (1916)'s work. Since then, mereology has been explored and developed by many authors, including some pivotal figures like H. S. Leonard, N. Goodman, W.V.O. Quine and D. Lewis. Roughly, the mereologists working in the analytic tradition aim at developing a formal theory which properly captures the genuine notion of parthood. They devise formal systems with different sets of axioms and definitions to constrain and regiment the behaviour of < in various ways, and try to understand which of these systems is the correct (or at least the best) theory of parthood relation. Such an evaluation depends essentially on the formal features of <, that is on the acceptability of the principles the theory at stake validates – i.e. its theorems.

Now, the reason to go into mereology is that parthood relation is one of the key resources of gluon theory, and some of the points I am going to make further ahead will rely on the specific features of the mereological theory Priest assumes. Then, in this section I am going to introduce the most important mereological notions, as well as present the most significant mereological theories available on the market. To do that, I will basically follow Varzi (2019), putting some emphasis on the aspects that

[20] To be precise, here we need the dual condition of **Act***: $w \Vdash^{-} \mathcal{A}B$ iff $@ \Vdash^{-} B$.

[21] Nonetheless, this issue may not be settled once and for all. For note that if we did not get in trouble with $\mathcal{A}$ it is because of **Act***. What I mean is that had we had a different truth condition, we could have been in trouble. Right now, I have no clue about that; but I am suggesting there might be a particular operator whose truth condition are such that what is true in a world affects what is true in another. And that could be problematic.

[22] Or relations, if you believe there are many different genuine notions of parthood relation – e.g. spatial parthood, functional parthood, and so on. Here we will assume that there is only one legitimate parthood relation: <.

are relevant for the assessment of gluon theory. But sometimes I will also refer to Cotnoir and Varzi (2021), which is arguably the best up-to-date and comprehensive publication about mereology available nowadays.[23] Before we begin, it should be stressed that there is no uncontentious principle in mereology. For it is relatively easy to find arguments against both the more substantive and the more basic features of $<$, since they often appear to betray our intuition. But for most of the mereological principles we are going to introduce I won't discuss such controversies. My primary aim is just to focus on the mereological issues related to gluon theory.

In the first place, we assume there is only one genuine and ontologically neutral parthood relation, $<$. That is, $<$ works the same no matter the nature of the entities – whether they be abstract, concrete, etc. Also, for the moment we bar the complications arising from intensional factors and dialetheism. Thus, we assume a standard first-order language with identity, supplied with the distinguished binary predicate constant $<$, and we take the underlying logic to be the classical predicate calculus with identity. Then, for a start we assume three principles that are simply meant to fix the intended meaning of «part»:[24]

$$x < x \tag{Reflexivity}$$
$$(x < y \land y < x) \rightarrow x = y \tag{Antisymmetry}$$
$$(x < y \land y < z) \rightarrow x < z \tag{Transitivity}$$

Reflexivity is certainly counter-intuitive: normally, we wouldn't say that everything is part of itself. But we shouldn't be worried about this misgiving. For the more intuitive relation of proper parthood, $\ll$, can be defined in terms of the (improper) parthood relation, $<$, and it is not difficult to prove that every mereological theory based on $<$ as a primitive can also be axiomatized by choosing $\ll$ in its place.[25] Antisymmetry guarantees that being mutual parts is a sufficient condition for identity. Transitivity enables us to consider the parts of a part of something as just parts of that something. These three principles together make $<$ a partial order and if we take them as the only three axioms (in addition to those of classical predicate calculus with identity) we get the theory called *Core Mereology* – **M** for short –, which represents the common starting point of all standard mereological theories.

What we can do now is to extend the set of axioms by adding further principles. There are many of them and they are usually categorised in two groups: composition principles and decomposition principles, depending on whether they take us from a

[23] To mention just two other bibliographical references, a very remarkable old-fashioned inquiry on mereology is Simons (1987), whereas for a good updated and mostly-philosophical introduction see Lando (2017).

[24] For ease of understanding, I switch back to using the standard quantifiers here. I will make use of $\mathfrak{S}$ and $\mathfrak{A}$ again as soon as we enter gluon theory. Note that all the principles I state here are universally closed, despite the universal quantification of their free variables is omitted for convenience of reading.

[25] This is exactly what Lesniewski (1916) did. See Lejewski (1957) for some considerations on that.

whole to its parts, or in the opposite direction. But before exploring new axioms, it is very useful to define at least some additional mereological predicates:[26]

Proper parthood:[27]	$x \ll y$	$\stackrel{\text{def}}{=}$	$x < y \land x \neq y$
Overlap:	$x \between y$	$\stackrel{\text{def}}{=}$	$\exists z\,(z < x \land z < y)$
Disjointness:	$x \not\between y$	$\stackrel{\text{def}}{=}$	$\neg(x \between y)$
Underlap:	$x \mathrel{\ddot{\bigcirc}} y$	$\stackrel{\text{def}}{=}$	$\exists z\,(x < z \land y < z)$

Note that by definition and pure logic, the proper parthood relation turns out to be irreflexive,[28] asymmetric[29] and transitive: i.e. $\ll$ is a strict partial order. Note also that if one assumed the existence of a universal entity – i.e. something of which everything is part – $x \mathrel{\ddot{\bigcirc}} y$ would be bound to hold for any x and y. Conversely, $x \between y$ would always hold if one assumed the existence of a null entity that is part of everything.

Let us now extend **M** in the direction of decomposition. Usually, the first idea to be formalized is that if something has a proper part, then it is not the only proper part it has – i.e. there is a reminder, something which supplements the proper part. There are a number of different ways to capture such an intuition, each of which corresponds to a specific 'supplementation' principle. Arguably, the principles of Weak and Strong Supplementation are the most important ones:

$$x \ll y \rightarrow \exists z\,(z < y \land z \not\between x) \qquad \text{(Weak Supplementation)}$$
$$y \not< x \rightarrow \exists z\,(z < y \land z \not\between x) \qquad \text{(Strong Supplementation)}$$

Weak Supplementation says that every proper part is supplemented by another disjoint part. It is quite a good principle to capture the previous idea, but many mereologists do not find it strong enough. This is the reason why they opt for Strong Supplementation, which says that if an object fails to include another among its parts, then there must be a remainder. Now, in the light of what we are going to say in the next sections of this chapter, it is crucial to examine the entailment relations

[26] In the literature, a number of different symbols are available to represent these notions. I've decided to create my own LaTeX mereological notation, though. This is designed to be reader-friendly, since it is based on the intuitive model which interprets parthood relation as spatial inclusion. I hope my choice will prove useful and won't lead to confusion – as it happens sometimes when there is not a unique notation.

[27] This is not the only definition we might choose for proper parthood. For at least one different definition is available: $x \ll y \stackrel{\text{def}}{=} x < y \land y \not< x$. These two definitions are equivalent in classical mereology (see below), but in non-classical mereologies they can have different impacts on the system.

[28] That is: $\neg(x \ll x)$, i.e. $x \not\ll x$.

[29] That is: $x \ll y \rightarrow y \not\ll x$.

among the principles we have introduced so far. If we extend **M** by adding Weak Supplementation to the set of axioms, we get the theory called *Minimal Mereology*, **MM**. But note that «[Weak] Supplementation turns out to entail Antisymmetry so long as parthood is transitive and reflexive: if x and y were proper parts of each other, contrary to [Antisymmetry], then every z that is part of one would also be part of–hence overlap–the other, contrary to [Weak Supplementation]» (Varzi, 2019, §3.1). Such an entailment is worth emphasizing, for it explains why Weak Supplementation tends to be explicitly rejected by those who do not endorse Antisymmetry and it will prove very important for a further discussion. On the other hand, if we extend **M** by adding Strong Supplementation to the set of axioms we obtain a theory called *Extensional Mereology*, **EM**, which is also an extension of **MM**. For together with Antisymmetry, Strong Supplementation entails Weak Supplementation, whereas the converse does not hold – i.e. there are **EM**-models that are **MM**-countermodels.

EM is highly controversial, though. For it validates a theorem called Extensionality which says that no composite objects with the same proper parts can be distinguished, and this is very much contentious:

$$\exists z\,(z \ll x) \lor \exists z\,(z \ll y) \to (x = y \leftrightarrow \forall z\,(z \ll x \leftrightarrow z \ll y)) \qquad \text{(Extensionality)}$$

The complaints against Extensionality can be divided into two categories. On the one hand, there are some arguments which challenge the left-to-right conditional in the consequent of the principle – i.e. $x = y \to \forall z\,(z \ll x \leftrightarrow z \ll y)$, which says that sameness of proper parts is necessary for identity. For we might find two identical objects with different proper parts. That is, there might be cases where identity survives mereological changes. Famous examples include the ship of Theseus, the cat Tibbles before and after losing its tail, and others, but they are relatively easy to tackle.[30] On the other hand, there are several arguments that question the right-to-left conditional in the consequent – i.e. $\forall z\,(z \ll x \leftrightarrow z \ll y) \to x = y$, which says that sameness of proper parts is sufficient for identity. This sufficiency feature of mereological extensionality is harder to address. For there are many alleged counterexamples to it, that is numerically different objects with exactly the same proper parts. For instance, the two words «else» and «seel», a nice bunch and a scattered bundle made of the same flowers, a statue and the lump of clay that constitutes it, and so forth. Even though controversial, more than one approach have been proposed to oppose each of them. But the debate is still open.

[30]For note that $x = y \to \forall z\,(z \ll x \leftrightarrow z \ll y)$ is just an instance of the Indiscernibility of Identicals: $x = y \to (\Phi x \leftrightarrow \Phi y)$, for any condition Φ. Then, whatever the solution for the general problem of identity thorough change, it will apply to the specific mereological issue at stake. For example, the problem vanishes if we consider the objects to be four-dimensional entities whose parts may extend in time as well as in space, or if identity is taken to be a contingent relation that may hold at some times or worlds, but not at others.

Before we proceed to examine some composition principles, there is one last couple of decomposition principles concerning atomism that is worth discussing. Let's assume the following definition for atom:

Atom: $Ax \stackrel{\text{def}}{=\!=} \neg\exists y\,(y \ll x)$

With that, we can state two substantive (and incompatible) mereological thesis:

$$\exists y\,(Ay \wedge y < x) \qquad \text{(Atomicity)}$$
$$\exists y\,(y \ll x) \qquad \text{(Atomlessness)}$$

Atomicity says that everything has atomic parts, and you just need to add it to a particular set of axioms to get the atomistic version of the corresponding mereological theory. Moreover, in conjunction with Reflexivity and Antisymmetry, Atomicity implies that everything is ultimately composed of atoms. But note that this is not equivalent to say that there is a mereological fundamental level, i.e. an ontological ground composed only by atoms that you eventually reach by descending the mereological structure of any object. For that you need to go in the direction of 'superatomism',[31] but we need not to face this matter here. On the contrary, Atomlessness says that every object has at least one proper part, and you just need to add it to a particular set of axioms to get the atomless version of the corresponding mereological theory. Together with Antisymmetry, Atomlessness captures the idea of a gunky world, a 'gunk' being an object that divides forever into smaller and smaller parts. Finally, a middle way between Atomicity and Atomlessness can also be held: there are atoms, though not everything need be made up of atoms; or there are atomless gunks, though not everything need be gunky.

Let's now focus on composition principles. The general strategy here is to impose some form of closure with respect to a specific mereological operation. To this regard, we can introduce the following notion of mereological sum (or fusion, or composition):[32]

Sum: $Szxy \stackrel{\text{def}}{=\!=} \forall w\,(z \circ w \leftrightarrow (w \circ x \vee w \circ y))$

where $Szxy$ is to be read as «z is a sum of x and y». Thus, a mereological sum of two objects, x and y, is an object z that overlaps exactly those things that overlap either x

[31] See Cotnoir (2013).

[32] This is not the only way we can define mereological sum. For different definitions see Varzi (2019, §4.2).

or y. Therefore, it is their minimal underlapper. Furthermore, we can even generalize such a notion to infinitary mereological sums. To do that there are various strategies. Later on we will introduce plural variables and plural quantification, but for now we can follow Varzi (2019, §4.3) and deploy sets of objects identified by predicates or open formulas. For «[s]ince an ordinary first-order language has a denumerable supply of open formulas, at most denumerably many sets (in any given domain) can be specified in this way» (Varzi, 2019, §4.3), and this is enough for our purposes. Thus:[33]

$$\text{General Sum:} \qquad Sz\Phi x \stackrel{\text{def}}{=\!=} \forall w\,(z \oplus w \leftrightarrow \exists x\,(\Phi x \wedge w \oplus x))$$

where «$Sz\Phi x$» may be read as «z is a sum of every x such that Φx». Then, z is the sum of the Φ-ers iff it overlaps all and only those things that overlap at least one Φ-er. Moreover, in the extensional contexts the sum of two (or more) objects is provably unique. Thus, we are allowed to use the definite description operator, ι, to define the (binary) Sum operator and its infinitary version for strong supplemented context:

$$\text{Sum operator:} \qquad x + y \stackrel{\text{def}}{=\!=} \iota z\, Szxy$$

$$\text{General Sum operator:} \qquad \bigoplus x\, \Phi x \stackrel{\text{def}}{=\!=} \iota z\, Sz\Phi x$$

Furthermore, for any mereological theory that validates Extensionality, such as **EM**, the Sum operator has all the basic Boolean properties – e.g. it is idempotent, commutative and associative.

Similarly, we can define a mereological product and its infinitary version. We can say that z is the product of the objects x and y (of the Φ-ers) iff all and only its parts are also parts of both x and y (of all the Φ-ers) – i.e. z is their maximal overlapper:

$$\text{Product:} \qquad Pzxy \stackrel{\text{def}}{=\!=} \forall w\,(w < z \leftrightarrow (w < x \wedge w < y))$$

$$\text{General Product:} \qquad Pz\Phi x \stackrel{\text{def}}{=\!=} \forall w\,(w < z \leftrightarrow \forall x\,(\Phi x \rightarrow w < x))$$

As before, in the extensional contexts we can define the (binary) Product operator and its infinitary version as follows:

$$\text{Product operator:} \qquad x \times y \stackrel{\text{def}}{=\!=} \iota z\, Pzxy$$

$$\text{General Product operator:} \qquad \bigotimes x\, \Phi x \stackrel{\text{def}}{=\!=} \iota z\, Pz\Phi x$$

[33]Here and below z and v are assumed not to occur free in Φ.

We are now ready to add a further (composition) principle that we can include into the set of axioms. The idea here is to impose that if some objects (the Φ-ers) satisfy a specific condition (say Ψ), then they have a mereological sum – i.e. there is an object which qualifies as their sum. Then:

$$\exists x \, \Phi x \wedge \forall x \, (\Phi x \rightarrow \Psi x) \rightarrow \exists z \, Sz\Phi x \qquad (\Psi\text{−Composition})$$

where the first conjunct in the antecedent is to guarantee that Φ picks out a non-empty set, and z is assumed not to occur free in Ψ. With regard to Ψ−Composition, the fundamental issue is to clarify what the condition Ψ is.[34] But this proved to be a tough question to answer. For instance, some have proposed to take causal connection as sufficient for fusing entities, while others the constituting of a life by the collective activity of the parts.[35] But there is no consensus, since any identification of Ψ carries some problems. Indeed, there is more: many mereologists opt for a complete liberalization of Ψ−Composition – i.e. a view called *mereological universalism*.[36] That is, they claim that any collection of objects whatsoever deserves a fusion: there would be no restriction to the entities you can sum.[37] Then, there would be the fusion of detached objects like the Moon and the Andromeda galaxy, but also that of fictional objects like Pegasus and Sherlock Holmes and, presumably, that of trans-categorical objects like π and Graham Priest.[38] This view about composition amounts to take Ψ to be a universal condition – i.e. a condition which is satisfied by the totality of objects. Then, we can free the composition principle from Ψ to get what is known as Unrestricted Composition:

$$\exists x \, \Phi x \rightarrow \exists z \, Sz\Phi x \qquad \text{(Unrestricted Composition)}$$

The theory obtained by adding the Unrestricted Composition axiom schema to **EM** is called *General Extensional Mereology*, **GEM**.[39] It is a very remarkable as well as

[34] This is known as van Inwagen's *Special Composition Question* (Van Inwagen and Van Inwagen, 1990). To be more precise, van Inwagen's question asks both the necessary and sufficient conditions for a collection of entities to have a mereological sum. On the contrary, Ψ is only sufficient for fusion.

[35] This is the view known as *organicism* and it was notably defended by Van Inwagen and Van Inwagen (1990).

[36] See e.g. Lewis (1991), Hawthorne (2006), Sider et al. (2001) and Cotnoir (2016).

[37] For the sake of completeness, the opposite view is also possible: there is no such a condition Ψ. That is, mereological sum is not possible. There are only atoms: no composite object inhabits our world. This view is called *mereological nihilism*. See for example Rosen and Dorr (2002).

[38] To be precise, this is quite wrong. For Unrestricted composition is silent about what kind of objects inhabit our domain. Depending on the ontology you choose, and given the generality of mereology, you get different situations. For instance, if we go nominalistic, we avoid such strange objects as trans-categorical fusions.

[39] Sometimes also known as *Classical Extensional Mereology*, **CEM**, or just *classical mereology*.

powerful theory, and it represents the mereological theory *par excellence*.

There are various equivalent axiomatizations for **GEM**. An elegant and simple way to characterize **GEM** is by noting that we can put Strong Supplementation and Unrestricted Composition together by using the General Sum operator, so to obtain a single substantial axiom schema. The principle we get is known as Unique Unrestricted Composition:

$$\exists x\, \Phi x \rightarrow \exists z\, (z = \bigoplus x\, \Phi x) \qquad \text{(Unique Unrestricted Composition)}$$

Given **M**, Unique Unrestricted Composition entails both Strong Supplementation and Unrestricted Composition. Thus, we just need the following minimal set of axioms (or axiom schemas) for **GEM**: Reflexivity, Antisymmetry, Transitivity and Unique Unrestricted Composition. However, the arguably most common and elegant axiomatization of **GEM** is that which relies on Transitivity, Unrestricted Composition and a principle called Uniqueness of Composition:

$$(x = \bigoplus z\, \Phi z \wedge y = \bigoplus z\, \Phi z) \rightarrow x = y \qquad \text{(Uniqueness of Composition)}$$

for every condition Φ. This principle guarantees that the mereological sum of any given collection of objects is unique, which is a fundamental requirement to cash out Strong Supplementation and, consequently, Extensionality. Also, note that in **GEM** we just need to rely on the mereological sum to define the product and other useful notions:

Product:	$x \times y$	$\stackrel{\text{def}}{=}$	$\bigoplus z\, (z < x \wedge z < y)$
Difference:	$x - y$	$\stackrel{\text{def}}{=}$	$\bigoplus z\, (z < x \wedge z \not\circ y)$
Complement:	$\overline{x}$	$\stackrel{\text{def}}{=}$	$\bigoplus z\, (z \not\circ x)$
Universe:	$\mathbf{u}$	$\stackrel{\text{def}}{=}$	$\bigoplus z\, (z < z)$

The power of **GEM** can be appreciated by considering that «it is isomorphic to the inclusion relation restricted to the set of all non-empty subsets of a given set, which is to say a complete Boolean algebra with the zero element removed» (Varzi, 2019, §4.4) – i.e. all the models of **GEM** are *complete join semilattices* and, conversely, all the *complete join semilattices* are model of **GEM**.[40] Moreover, it can be shown that if we extend **GEM** by imposing the existence of a null entity – i.e. $\exists x \forall y\, (x < y)$ – plus

[40] This general result is owed to Tarski (1935).

some adjustments to both Strong Supplementation and Unrestricted Composition, we get a full Boolean algebra. And «[t]his shows that, mathematically, mereology does indeed have all the resources to stand as a robust and yet nominalistically acceptable alternative to set theory, the real source of difference being the attitude towards the nature of singletons» (Varzi, 2019, §4.4).

There are several other mereological theories in the literature. This is not the place to go into details and assess each one of them. Nonetheless, one particular class of such theories is worth to be discussed here: non-well-founded mereologies. For Priest's favoured mereology for gluon theory belongs to this family.

3.2.1 Non-well-founded mereologies

Likewise set theories, a mereological theory is said to be well-founded if every nonempty subset of the domain has a $\ll$-minimal element, whose definition is as follows: x is a $\ll$-minimal element of a set S if there is no y in S such that $y \ll x$.[41] Thus, a non-well-founded mereology (NWFM) is a mereological theory that is not well-founded: there is at least one nonempty subset of the domain with no $\ll$-minimal element.

Given a set of objects, there are essentially two ways we can abolish the $\ll$-minimal element. The first strategy is to allow for an infinite descending $\ll$-chain, that is to include at least one infinitely decomposable object in the domain – i.e. a gunk. For ease of understanding, we can use simple graphs to represent some models which realize non-well-foundedness, where nodes represent objects and arrows represent proper parthood relations (if the arrow goes from a to b, read $a \ll b$). As usual, dotted edges mean that the structure keeps on developing indefinitely in the very same way. Thus, the models of figure 3.1 (left and center) realize such a first strategy. Note that they are unsupplemented models, since they violate Weak Supplementation and any other stronger supplementation principle. But the violation of supplementation is not necessary for non-well-foundedness; for we may have also infinite descending supplemented $\ll$-chain, as shown in figure 3.1 (right). This first approach requires a countably infinite set of objects and is not incompatible with Atomicity, as proved again by the graph of figure 3.1 (right).

The second option is to allow for loops, that is closed $\ll$-chains. Figure 3.2 shows some simple loopy models. Such loops may be composed by just one object – a mono-loop, sometimes referred to as *Quine atom* –, so that this is a proper part of itself ($a \ll a$); or two objects, which are mutual proper parts, $a \ll b$ and $b \ll a$ – a

[41] Note that well-foundedness is defined in terms of proper parthood – and not improper parthood. The reason is very simple: since $<$ is meant to include identity, and therefore is usually taken to be reflexive, for every nonempty subset of the domain there would always be a $<$-minimal element. Thus, every mereological theory would be non-well-founded, unless we are willing to give up Reflexivity for $<$. Then, to properly capture the feature of non-well-foundedness we need to exclude identity, i.e. we have to use $\ll$.

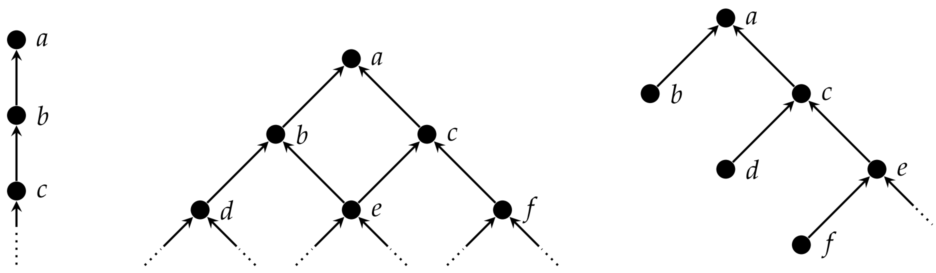


FIGURE 3.1: Some infinitely descending non-well-founded models. The left and central models violate Weak and stronger Supplementation principles, whereas the right model is supplemented (and atomistic).

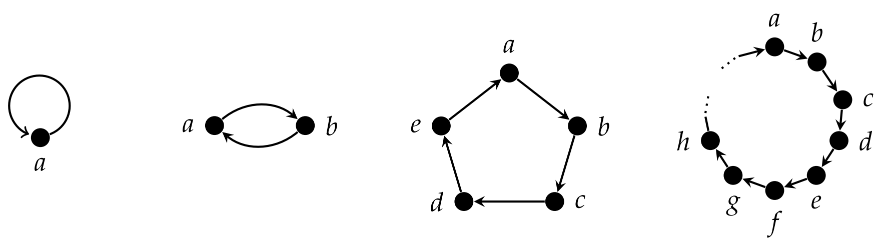


FIGURE 3.2: Some non-well-founded loopy models. They all are unsupplemented (i.e. violate Weak Supplementation), even if this is not necessary to infringe well-foundedness. From left to right: a mono-loop (a self-proper part), a bi-loop (two mutual proper parts), a penta-loop and an infinite-loop.

bi-loop. But they may also be composed by more objects, up to an infinite number of them.

There may be various reasons to go into non-well-foundedness.[42] For instance, non-well-founded mereologies have been invoked to account for some exotic objects, such as Burges's Aleph and the Trinity; or to solve some puzzles of composition, such as that which concerns Tibbels the cat; or, again, to make sense of some curious scenarios related to time travel. But the reason why I am interested in this family of mereological theories is precisely because the mereology Priest assumes for gluon theory belongs to it. More specifically, we will focus on NWFMs with loopy models only. This too is a broad group of theories, since several kinds of loopy models satisfying different principles are conceivable. Cotnoir and Bacon (2012) have examined mereologies corresponding to mono-loops and bi-loops (but that also extend to larger loops), and we will start by giving a look at their analysis.[43] Though, the kind of NWFM Priest strives for is more demanding. But arguably, as I will show in §3.5, he has to leave some *desideratum* behind to make sense of the mereological structure of gluon theory.

[42]For a brief survey, see Cotnoir and Bacon (2012, §1).

[43]There might be other NWFMs than that of Cotnoir and Bacon (2012). Here, I will not try developing such alternatives. Later in the chapter, I will take their NWFM as a reference, and compare it with Priest's preferred mereology for gluon theory to see whether or not the latter meets the former.

How then a mereological theory is required to be in order to have both mono-loopy and bi-loopy models? Well, it depends. First of all, it is worth mentioning a very important warning highlighted by Cotnoir and Varzi (2021, p. 68): «definitions that are equivalent in classical mereology may fail to be so in other, weaker theories. When considering non-classical axiom systems, one should therefore pay careful attention to definitional matters». Thus, let us here make explicit Cotnoir and Bacon (2012)'s initial assumptions. They take proper parthood ($\ll$) as primitive, and improper parthood ($<$), overlap ($\between$) and mereological sum (S) to be defined as follows:

$$\text{Parthood relation:} \quad x < y \overset{\text{def}}{=\!=} x \ll y \lor x = y$$

$$\text{Overlap:} \quad x \between y \overset{\text{def}}{=\!=} \exists z\,(z < x \land z < y)$$

$$\text{Sum:} \quad Sz\Phi x \overset{\text{def}}{=\!=} \forall w\,(z \between w \leftrightarrow \exists x\,(\Phi x \land w \between x))$$

Then, the axioms they opt for are:[44]

$$x \ll y \land y \ll z \to x \ll z \qquad (\ll\text{-Transitivity})$$

$$y \not< x \to \exists z\,(z < y \land z \not\between x) \qquad (\text{Strong Supplementation})$$

$$\exists x\,\Phi x \to \exists z\,Sz\Phi x \qquad (\text{Unrestricted Composition})$$

Now, let us ignore Unrestricted Composition for a moment. Since $\ll$ lacks both Irreflexivity and Asymmetry, we may have $x \ll x$ (mono-loops) and $x \ll y \land y \ll x$ (bi-loops), for some x and y. Note that $\ll$-Transitivity implies that if two distinct objects are mutual proper parts, they are also self-proper parts. That is, the bi-loopy model of figure 3.2 is ruled out, and only bi-loopy models such as that in figure 3.3 are allowed.

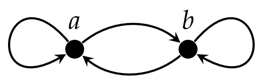


FIGURE 3.3: In Cotnoir and Bacon (2012)'s NWFM mutual proper parts are always self-proper parts.

[44] To be precise, Cotnoir and Bacon (2012) show that two equivalent axiomatizations for their NWFM are possible, the other one relying on a different definition of fusion and the axiom of Complementation in the place of Strong Supplementation.

$<$ inherits Reflexivity from $=$ and Transitivity from $\ll$ and $=$, but it is not antisymmetric since two objects can be mutual parts without being identical. Moreover, given $\ll$-Transitivity and the presence of distinct mutual parts, we lose Extensionality. To see that, consider the model of figure 3.4 (but also that of figure 3.3). Here, $\ll$-Transitivity and Strong Supplementation hold and a and b are distinct objects, even though they have all and only the same proper parts. Therefore, such a model is a countermodel for Extensionality. (And that should not be surprising since Transitivity plus Extensionality rules out loops.) Note that this model includes two mereological atoms, c and d, and validates Atomicity: every object has at least one atom as improper part. Thus, Atomicity is compatible also with loopy models.

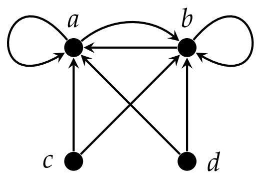


FIGURE 3.4: A model of Cotnoir and Bacon (2012)'s NWFM which contradicts Extensionality and validates Atomicity.

Strong Supplementation is the natural choice to make in terms of decomposition principles. For its weaker versions, such as Weak Supplementation, turn out to be incompatible with self-parts and distinct mutual parts. Moreover, in §3.2 we said that **M** plus Strong Supplementation implies Weak Supplementation. But such an implication crucially relies on Antisymmetry of parthood, and therefore Cotnoir and Bacon (2012)'s NWFM is not weakly supplemented even though it is strongly supplemented.

Let us now examine how Cotnoir and Bacon (2012)'s NWFM behaves with loops larger than bi-loops. Clearly, both the penta- and the infinite-loops of figure 3.2 are not models of such a mereology. For they contradict $\ll$-Transitivity, as well as Strong Supplementation. Instead, a penta-loopy model is that of figure 3.5:

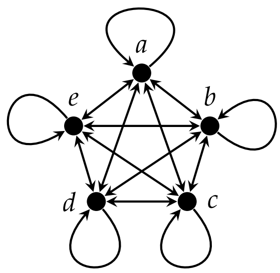


FIGURE 3.5: Penta-loopy model for Cotnoir and Bacon (2012)'s NWFM. Note the double sided arrows I have used for ease of reading the graph.

Basically, as soon as we close a ≪-chain – i.e. we get a loop – ≪-Transitivity imposes to every pair of objects in the loop to be mutual proper parts, and it forces every object in the loop to be self-proper part as well. Further, Strong Supplementation holds in the model of figure 3.5 because its antecedent is false, which makes the conditional true. Since every object is a proper part of every object (including itself) we can call such a kind of structures *totally interpenetrating models*. These always validate Atomlessness.

Now, take a totally interpenetrating model. Because of ≪-transitivity, as soon as a new object becomes a proper part of one of the objects in the loop it also becomes proper part of every other object. But then, Strong Supplementation fails. To see that, consider the graph of figure 3.6 (left). Let $x = c$ and $y = a$, so that the antecedent of the principle is true. But the consequent is false, since there is no object that is part of a and does not overlap c. For in the model every object overlap c. What we need then is another object which is also proper part of every object in the loop, but that does not overlap with c. That is, we need a model like that of figure 3.4. Moreover, this move proves to be effective whatever the number of the objects the loop is made of. See e.g. figure 3.7 for the tri-loopy version of this kind of model. And since not every object of these models is a proper part of every object, we can call such a group of structures *partially interpenetrating models* for Cotnoir and Bacon (2012)'s NWFM.

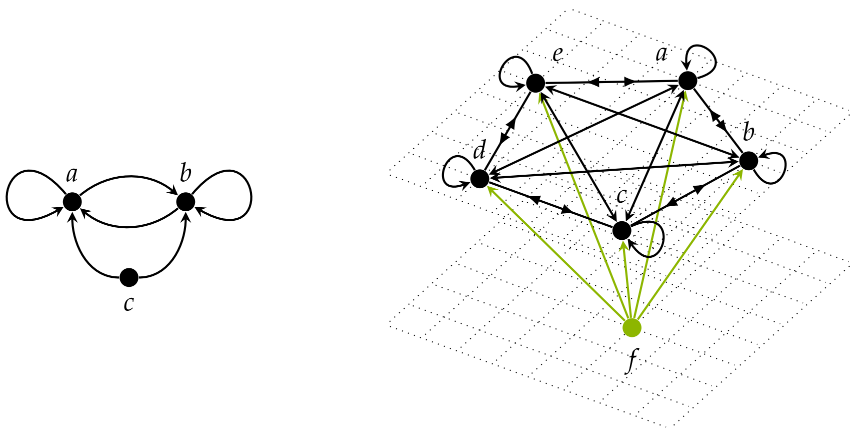


FIGURE 3.6: Two example of forbidden structures for Cotnoir and Bacon (2012)'s NWFM. They violate Strong Supplementation in the very same way, the only difference being that in the left structure we have a bi-loop, whereas in the right structure we have a penta-loop. Note that in the right graph I have deployed a 3D-graph and colors to make the figure more comprehensible.

So much for decomposition. What happens if we consider Unrestricted Composition? Take the graph 3.3: is it a model for Cotnoir and Bacon (2012)'s NFWM? The answer is yes, since Unrestricted Composition is satisfied. To see that, take the extension of Φ to be $\{a, b\}$. Now, consider a (or, equivalently, b). Since it overlaps all and only a and b, it counts as their fusion. And since a and b are their own fusions –

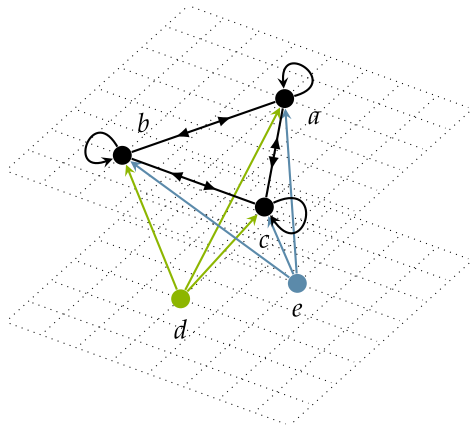


FIGURE 3.7: A partially interpenetrating model of Cotnoir and Bacon (2012)'s NWFM.

i.e. Saa and Sbb – we can conclude that the model satisfies Unrestricted Composition. In addition, from this model we can see another very important aspect of such a NWFM: Uniqueness of Composition does not hold. For a and b are two distinct fusions of the very same objects, i.e. themselves. Also, exactly the same features extend to all totally interpenetrating models. Consider e.g. the penta-loop of figure 3.5 and one of its objects, say a. Such an object turns out to be the fusion of $\{a, b, c, d, e\}$ and the same is for every object of the loop. Moreover, a is also the fusion of whatever combination of the remaining objects of the loop. That is, $Sa\Phi x$ for all the $2^5 - 1 = 31$ extensions of Φ we can get with 5 distinct objects. Thus, all totally interpenetrating models are indeed models of Cotnoir and Bacon (2012)'s NWFM.

But what about the partially interpenetrating models? Do they always satisfy Unrestricted Composition? As expected, the answer is no. Let us begin by considering the figure 3.4. As I show below, for every combination of objects we can fuse we always have its fusion(s) as one (or some) of the entities in the domain. Let $\pi(\Phi)$ represents the extension of Φ. Then, we have:

- if $\pi(\Phi) = \{a\}$ then $Sa\Phi x \wedge Sb\Phi x$
- if $\pi(\Phi) = \{b\}$ then $Sa\Phi x \wedge Sb\Phi x$
- if $\pi(\Phi) = \{c\}$ then $Sc\Phi x$
- if $\pi(\Phi) = \{d\}$ then $Sd\Phi x$
- if $\pi(\Phi) = \{a, b\}$ then $Sa\Phi x \wedge Sb\Phi x$
- if $\pi(\Phi) = \{a, c\}$ then $Sa\Phi x \wedge Sb\Phi x$
- if $\pi(\Phi) = \{a, d\}$ then $Sa\Phi x \wedge Sb\Phi x$
- if $\pi(\Phi) = \{b, c\}$ then $Sa\Phi x \wedge Sb\Phi x$
- if $\pi(\Phi) = \{b, d\}$ then $Sa\Phi x \wedge Sb\Phi x$

- if $\pi(\Phi) = \{c, d\}$ then $Sa\Phi x \wedge Sb\Phi x$
- if $\pi(\Phi) = \{a, b, c\}$ then $Sa\Phi x \wedge Sb\Phi x$
- if $\pi(\Phi) = \{a, b, d\}$ then $Sa\Phi x \wedge Sb\Phi x$
- if $\pi(\Phi) = \{a, c, d\}$ then $Sa\Phi x \wedge Sb\Phi x$
- if $\pi(\Phi) = \{b, c, d\}$ then $Sa\Phi x \wedge Sb\Phi x$
- if $\pi(\Phi) = \{a, b, c, d\}$ then $Sa\Phi x \wedge Sb\Phi x$

Thus, the graph of figure 3.4 satisfies Unrestricted Composition and is therefore a model of Cotnoir and Bacon (2012)'s NWFM. But as soon as we introduce a third object with no proper parts – as it is e in the graph of figure 3.8 –, Unrestricted Composition is violated. To see that, consider the fusion of c and d. None of the objects of the graph is such a fusion. For both a and b cannot be that fusion since they overlap e, which does not overlap neither c nor d. And of course, c, d and e are not this fusion either. Thus, this is not a model for Cotnoir and Bacon (2012)'s NWFM.

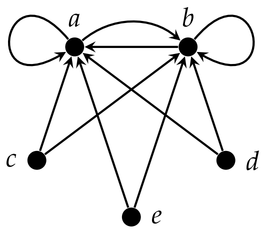


FIGURE 3.8: A structure which validates both $\ll$-Transitivity and Strong Supplementation, but that invalidates Unrestricted Composition.

If we want to fix the graph of figure 3.8 to get a model of Cotnoir and Bacon (2012)'s NWFM, what we need is a minimal upper bound for each subset of the object of the domain. That is, we have to introduce three more objects which are the fusions of the following subsets of entities: $\{c, e\}$, $\{e, d\}$ and $\{c, d\}$. Thus, we get the (atomistic) model of figure 3.9. Moreover, the same considerations apply to every similar model, whatever the number of the objects the loop is made of.

Proceeding our analysis, next step is to go in the opposite direction: instead of looking at loops the objects of which have proper parts outside the loop, now we want to examine what happens if the objects in the loop are proper parts of another object outside the loop. For a start, consider an object which has one object of the loop as proper part. Because of $\ll$-Transitivity, it gains all the remaining objects of the loop as proper parts as well. But such a scenario, which is represented in figure 3.10 (left) invalidates Strong Supplementation. For $c \not< a$, but there is no z such that

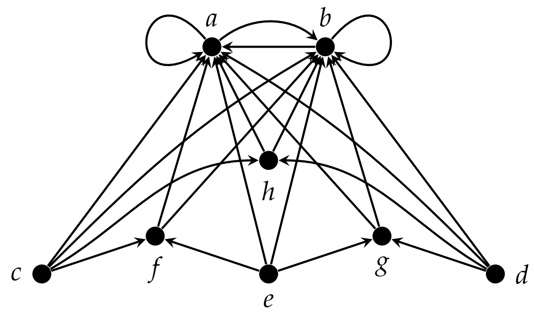


FIGURE 3.9: A model of Cotnoir and Bacon (2012)'s NWFM.

$(z < c \land z \not\circ a)$. To fix this problem, we can simply add another object outside the loop, d, such that it is a proper part of c. Moreover, Unrestricted Composition holds in this model: for any combination of objects there is their fusion. Thus, figure 3.10 (right) shows a model of Cotnoir and Bacon (2012)'s NWFM – notably, a partially interpenetrating model. Also, note that we can have a whole loop in the place of d (figure 3.11, left), but that we cannot have more than one (immediate) proper part of c outside the loop (be that a single object or a whole loop), because of Unrestricted Composition (figure 3.11, right).

FIGURE 3.10: Left: A structure that invalidates Strong Supplementation. Right: a model of Cotnoir and Bacon (2012)'s NWFM where the whole loop is a proper part of an external object – that is, every object of the loop is a proper part of an external object.

Of course, we can keep on extending the models by adding new objects and loops, so to get more complicated scenarios. But we can just stop our analysis here. Though, there is still one feature of Cotnoir and Bacon (2012)'s NFWM that needs to be discussed: is the presence of a null object and a universal object compatible with such a mereological theory?

We have already met both the null and the universal objects in §3.2. A null object is an entity which is part of everything, whereas a universal object is an entity everything is part of. Assuming their existence corresponds to accept the following postulates:

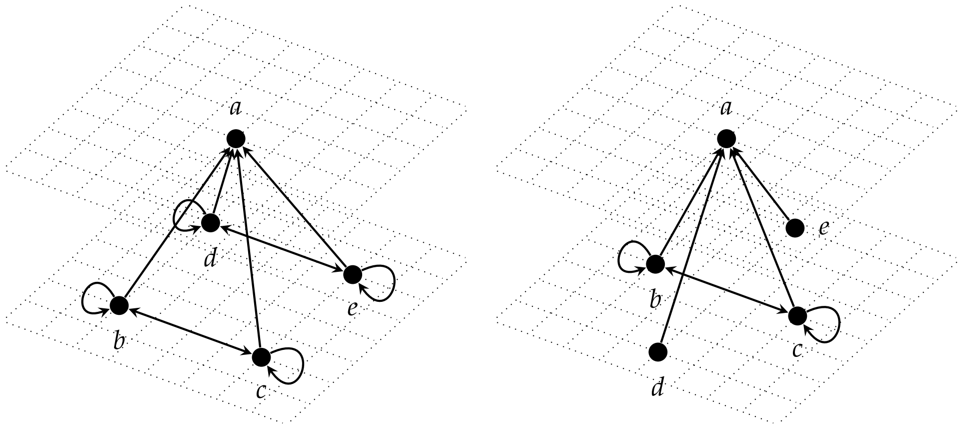


FIGURE 3.11: Left: a model of Cotnoir and Bacon (2012)'s NWFM. Right: a structure that invalidates Unrestricted Composition. For there is no fusion of d and e.

$$\exists x \forall y \ (x < y) \tag{Bottom}$$

$$\exists x \forall y \ (y < x) \tag{Top}$$

In **GEM**, the existence of a universal object is guaranteed – i.e. Top is a theorem – and that turns out to be unique (so that we can call it the universe, **u**). On the contrary, postulating the null object in **GEM** brings a bad consequence: a model for such a theory must have one single object. That is, **GEM** plus Bottom yields a monistic ontology, as testified by the following theorem:

$$\exists x \forall y \ (x = y) \tag{Oneness}$$

For «given [Bottom], the Antisymmetry axiom [...] will immediately entail that the atom in question is unique, while the Reflexivity axiom [...] will entail that it overlaps everything, hence that everything overlaps everything. This means that under such axioms the [Weak] Supplementation principle [...] cannot be satisfied except in models whose domain includes a single element» (Varzi, 2019, §3.4).[45] In other words, in **GEM** we can prove the following theorem:

$$\exists x \exists y \ (x \neq y) \rightarrow \neg \exists x \forall y \ (x < y) \tag{No Bottom}$$

[45] Actually, as we mentioned in §3.2 the existence of a null entity in **GEM** can be added, provided we make some adjustments to the definition of null object and also to both Strong Supplementation and Unrestricted Composition. For more on that, see Cotnoir and Varzi (2021, §4.5).

Then, what about Bottom and Top in Cotnoir and Bacon (2012)'s NWFM? Let us begin with Top. For a start, note that a universal object is compatible with such a mereology. We have it in all the models we have seen so far. For instance, in the model of figure 3.11 (left) a is the universal object. The difference with **GEM** is that the universal object may not be unique, since Uniqueness of Composition does not hold. For example, in the model of figure 3.9 both a and b are universal objects. But we can go further and say that at least one universal object is guaranteed also by Cotnoir and Bacon (2012)'s NWFM. For Unrestricted Composition ensures the existence of the fusion of every object, so that it has every object as part (including itself, since parthood is reflexive). Thus, such a fusion is a universal object. Moreover, a stricter notion of universal object becomes available in loopy context. We can define a *strict universal object* as an object which has every object (including itself) as proper part. Of course, this object is ruled out in **GEM** since loops are not allowed, but Cotnoir and Bacon (2012)'s NWFM is compatible with it – even if its existence is not ensured by a theorem. As examples, see the models of figures 3.12 and 3.7.

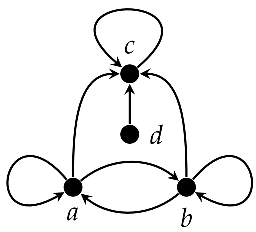


FIGURE 3.12: A model of Cotnoir and Bacon (2012)'s NWFM with a strict universal object, c.

With regard to the null object, the situation is a little more complicated. In the first place, we can observe that Bottom holds in any totally interpenetrating model. For here every entity is a null object, since every object is part of every object. But unless we are willing to accept Oneness, partially interpenetrating models are incompatible with the null object. To see that, recall that a partially interpenetrating model is a model where at least one object is not a proper part of another object. Thus, there is (at least) a pair of objects satisfying the antecedent of Strong Supplementation. However, its consequent cannot be satisfied. For a null object make every object overlap every object. Therefore, in this case a null object would cause a violation of Strong Supplementation. The moral is that Cotnoir and Bacon (2012)'s NWFM tolerates null objects, but only for totally interpenetrating models – as well as for models with one single object. But this is too fast. For we may question whether Bottom correctly captures the notion of "null object" in case of non-well-founded models. For a «null object is supposed to be a genuine mereological bottom: something that is part of everything *and* has no proper parts of its own» (Cotnoir and Varzi, 2021, p. 136). And

Bottom does not seem to do justice to this narrower notion. Then, in the contexts where Antysimmetry for parthood fails, we can go for a stricter postulate for the existence of a (more genuine) null object:

$$\exists x \forall y\ (x < y \land y \nless x) \qquad \text{(Strict Bottom)}$$

That is, a strict null object is characterised as something which is part of everything, and with no proper part – i.e. an atom. Consequently, Strict Bottom immediately rejects any totally interpenetrating model. In addition, such a principle is incompatible with partially interpenetrating models for exactly the same reasons Bottom is. Thus, Strict Bottom is more generally incompatible with Cotnoir and Bacon (2012)'s NWFM – unless we accept Oneness. But again, the matter is not settled yet. For we said that in the case of **GEM** we can make room for a (strict) null object by making some adjustments. The insight here is that the reason why a (strict) null object is incompatible with **GEM** is that its axioms were biased from the beginning, so not to allow for such an entity. Thus, we can rephrase them according to (Strict) Bottom and get a full Boolean algebra. Therefore, we may wonder: is this possible even for Cotnoir and Bacon (2012)'s NWFM? The answer is yes. For we can include Strict Bottom as one of the axioms of (this new version of) Cotnoir and Bacon (2012)'s NWFM, so that we can define the null object as follows:

$$\text{Strict null object:} \qquad \mathbf{n} \stackrel{\text{def}}{=} \iota x \forall y\ (x < y \land y \nless x)$$

We can speak of *the* null object because it is unique. To see that, assume two distinct null objects and use *reductio* on the *absurd* you can draw from that hypothesis. Then, following Cotnoir and Varzi (2021, p. 141), we define:

$$\text{Solid disjointness:} \qquad x \not\bigcirc_{\mathbf{n}} y \stackrel{\text{def}}{=} \neg\exists z\,(z < x \land z < y \land z \neq \mathbf{n})$$

That is, x is solidly disjoint from y iff there is nothing both of them overlap, except for the strict null object. Finally, we just need to replace Strong Supplementation with what we can call Solid Strong Supplementation:

$$y \nless x \rightarrow \exists z\,(z < y \land z \not\bigcirc_{\mathbf{n}} x) \qquad \text{(Solid Strong Supplementation)}$$

Thus, the strict null object is in a way compatible with Cotnoir and Varzi (2021)'s NWFM. But many changes have occurred and some comments are in order. First of all, note that the structures of figure 3.6 become models of such a new version of Cotnoir and Varzi (2021)'s NWFM, whereas e.g. those of figures 3.4, 3.5 and 3.10 (right) are not models anymore. Moreover, even if non-well-foundedness is preserved – since loops are still possible and therefore there can be at least one nonempty subset of the domain with no $\ll$-minimal element – it turns out to be severely limited. For totally interpenetrating models are banished and only partially interpenetrating atomistic models are allowed. In other words, despite deserving the title of non-well-founded mereology, such a new mereological theory is very distant in spirit from Cotnoir and Varzi (2021)'s NWFM.

Before concluding, a quick observation. We have seen that all totally interpenetrating models validate Atomlessness, and that there are some partially interpenetrating atomistic models. Then, we may wonder: are there also partially interpenetrating atomless models? The answer is yes. Consider the graph of figure 3.4. There are essentially two strategy to make this model atomless: developing a suitable infinite descending pattern below c and d, or developing an appropriate mereological structure for c and d which ends with at least one loop. Following this second strategy, I show a simple example of partially interpenetrating atomless model in figure 3.13. As an example of a partially interpenetrating model which does not satisfy neither Atomicity nor Atomlessness, see again the graph of figure 3.10 (right).

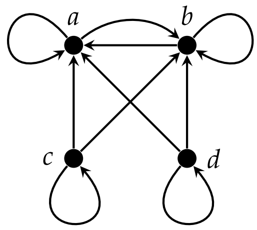


FIGURE 3.13: Another example of partially interpenetrating atomless model of Cotnoir and Bacon (2012)'s NWFM.

We should now have a sufficient understanding of both classical and non-well-founded mereologies. Together with well-foundedness, the main features we have to leave behind in passing from classical mereology to Cotnoir and Bacon (2012)'s NWFM are: Antisimmetry (for <), Weak Supplementation, Extensionality and Uniqueness of Composition. Then, provided that Priest's mereology for gluon theory takes the form of Cotnoir and Bacon (2012)'s NWFM – something we will deal with in §3.5 – a good question we can pose is whether it can really do without such mereological principles. But now, before moving to the next step – i.e. plural quantification – let us say something also about mereology in modal contexts.

3.2.2 Parthood in modal contexts

The powerful language of gluon theory allows to express modality – among other things. Thus, we may wonder how mereology – i.e. the parthood relation and its related notions – functions in this more complicated logical setting, based on a world semantic framework.[46] How does a mereological theory behave across possible and impossible worlds? How does parthood relation interact with the alethic modal operators of necessity ($\Box$) and possibility ($\Diamond$)?

The parthood relation is just a predicate governed by the axioms of the mereology we choose. In the same way we can have logically impossible worlds (i.e. worlds where logical laws are different that those holding in the actual and other possible worlds), metaphysically impossible worlds (i.e. worlds where metaphysical laws are different that those holding in the actual and other possible worlds) and physically impossible worlds (i.e. worlds where physical laws are different that those holding in the actual and other possible worlds), we can also have mereologically impossible worlds: worlds where the parthood relation behaves differently than the way it behaves in the actual and other mereologically possible worlds. Such worlds may validate different mereological theories (i.e. other sets of axioms), up to be completely anarchic with respect to parthood – there would be no regimentation of such a relation. Therefore, we might wonder how mereological possibility relates to metaphysical possibility. The matter is not settled in Priest (2014b), but here I want to quickly discuss how I see it. I tend to think of parthood as one of the fundamental relations holding between the entities inhabiting the ontology of our world. Thus, if we assume that metaphysical laws are the laws regimenting the behaviour of such fundamental relations, mereological laws constitute a subset of metaphysical laws.[47] Therefore, all metaphysically possible worlds are mereologically possible worlds, whereas the converse does not hold. And consequently, mereologically impossible worlds are *ipso facto* metaphysically impossible worlds. But note that this does not mean that all mereologically possible worlds have the very same mereological structure of the actual world. For satisfying the same mereological theory does not imply being exactly the same model of it. For instance, if $a < b$ in the actual world, then it does not follow that $a < b$ in every metaphysically possible world. That is a further stronger requirement, captured by the two following principles introduced by Chisholm (1973, p. 581), where the box is supposed to represent metaphysical necessity:

[46] Two useful references on mereology and modality are Uzquiano (2014) and Cotnoir and Varzi (2021, §6.2). Note that these only discuss the more orthodox world semantic framework, where only possible worlds are involved

[47] As an example, if you choose **GEM** as the mereology of the actual world, then all metaphysically possible worlds validate e.g. Extensionality.

$$\forall x \forall y\ (x < y \rightarrow \Box(\exists z(z = y) \rightarrow x < y)) \qquad \text{(Mereological Essentialism)}$$
$$\forall x \forall y\ (x < y \rightarrow \Box(\exists z(z = x) \rightarrow x < y)) \qquad \text{(Holological Essentialism)}$$

Mereological Essentialism says that if x is part of y, then it is metaphysically necessarily so: it is part of y in every metaphysically possible world in which y exists. Holological Essentialism states that if x is part of y, then it is metaphysically necessarily so: it is part of y in every metaphysically possible world in which x exists. Then, the two principles together fix the same common mereological structure for all metaphysically possible worlds. Now, Priest is silent about whether such principles hold in GT. However, I believe he would have none of them. For they make metaphysically possible worlds mereologically indistinguishable and he would have to accept that e.g. the actual parts of his body are all and only the parts of his body in every metaphysically possible world. But it is very easy to conceive a possible world where e.g. Priest has just one leg (sorry, Graham!). Also, given their relevance and impact, if he had accepted them he would have made it explicit. But there is no reference about these principles in Priest (2014b).

Of course, there is more to be said about the interaction between mereology and modality. But we can just stop here and move forward to the next topic: plural logic.

3.3 Plural logic

Plural quantification is a logical machinery developed by Boolos (1984, 1985) which accounts for the plural reference and quantification we normally use in natural languages. For Boolos suggests that «as the singular quantifiers [...] get their legitimacy from the fact that they represent certain quantificational devices in natural language, so do their plural counterparts» (Linnebo, 2017, §1). And it does seem that expressions such as «for any things» and «there are some things» are legitimate and do not need to be paraphrased away in terms of singular quantification – i.e. plural quantifiers might be reasonably admitted among our primitive logical notions. In addition, plural quantification proves to be an extremely powerful tool and some philosophers look at it as purely logical – that is universally applicable, ontologically innocent, and perfectly well understood. Be that as it may, Priest sometimes deploys plural quantification to discuss gluon theory, and this makes it worth to be presented here. For this purpose, I will essentially follow Linnebo (2017).

Let us see now how to get first-order plural logic (with identity), i.e. first-order logic (with identity) plus plural quantification.[48] The first step is to extend the starting formal language with the following ingredients:

Plural variables:	$xx, yy, zz, \dots$
Plural constants:	$aa, bb, cc, \dots$
The logical predicate:	$\prec$

where $\prec$ is a dyadic logical predicate to be thought of as the relation «is one of». Moreover, we allow for predicates with plural arguments of any adicity. This is crucial, since we want to account for both distributive and non-distributive (or collective) plural predication. For the record, a predicate P taking plural arguments is:

- *distributive* just in case it is analytic that P holds of some things xx iff P holds of each y such that $y \prec xx$;
- *non-distributive* iff P doesn't meet the distributivity condition we have just stated.

The syntax is also extended very intuitively. We want to include formulas such as $\exists xx\, \alpha$, when α is a formula, and $t \prec T$, when t is a singular term and T a plural term. Instead, identity remains unchanged and relates only singular terms. Thus, we get a more powerful language than that of first-order logic (with identity). For it can be shown that there are some sentences[49] of the plural language we have just introduced that cannot be expressed with the resources of first-order (singular) logic alone.

Next step is the construction of a suitable theory based on this plural language. In short, we can just add the following axioms (or axiom schemas) to the canonical first-order logic with identity:

$$\exists x\, \Phi x \to \exists yy\, \forall x\, (x \prec yy \leftrightarrow \Phi x) \qquad \text{(Plural Comprehension)}$$
$$\forall xx\, \exists y\, (y \prec xx) \qquad \text{(Non-Emptiness)}$$
$$\forall xx\, \forall yy\, (\forall z\, (z \prec xx \leftrightarrow z \prec yy) \to (\Phi xx \leftrightarrow \Phi yy)) \qquad \text{(Extensionality for Pluralities)}$$

Plural Comprehension says that if something is Φ then there are some things such that everything is one of them iff it is Φ. Non-Emptiness guarantees that all pluralities

[48] This is the plural logic we obtain if plural variables are constrained to range on pluralities of individuals only – and not on e.g. plurality of properties as with in second-order plural logics.

[49] A famous example is the so-called Geach-Kaplan sentence: «Some critics admire only one another».

are non-empty. Extensionality for Pluralities ensures that all coextensive pluralities are indiscernible. Following Linnebo (2017), we can call such a theory **PFO+**. And of course, this is just one example of a plural theory, since we can arguably implement the plural quantification machinery to any theory, at least in principle.

Plural quantification has many applications.[50] There's no need to discuss them here, but we can just note that it may be convenient to use plural quantification to rephrase some of the mereological principles we have introduced in §3.2. For instance:

$$(x = \bigoplus zz \wedge y = \bigoplus zz) \rightarrow x = y \qquad \text{(Uniqueness of Composition)}$$

$$\exists y\, (y = \bigoplus xx) \qquad \text{(Unrestricted Composition)}$$

where the seemingly unbound variables x, y, zz and xx are meant to be universally quantified. ($\bigoplus zz$ should be read as «the fusion of the zz».) Later this will prove useful to facilitate our examination of some consequences of Priest's preferred mereology for gluon theory.

To conclude this section, let me say a few words about what is known as the thesis of Composition as Identity (CAI). Unrestricted Composition does seem to generate an 'ontological explosion': since whatever collection of objects has a mereological sum, if we accept the existential Quinean dictate we get an extremely overpopulated and possibly queer ontology, also composed by an impressive number of transcategorical objects. And since ontological parsimony and intuitiveness are usually seen as theoretical virtues, we may want to avoid such an undesired inconvenience. Thus, it can be argued that a sum is nothing over and above its constituent parts: the sum is just the parts «taken together» (Baxter, 1988, p. 193). This amounts to say that composition is just a specific case of the identity relation – or, in other words, that the extension of the composition relation is a subset of the extension of the identity relation. For in this perspective, the composite object *is* its parts.

Formally, CAI can be stated resorting to plural logic:

$$\text{Composition as Identity:} \qquad Cxxy \overset{\text{def}}{=\!=} xx = y$$

Let's make a few comments. Firstly, it should be evident that the identity relation occurring in this definition is different from the usual identity relation. For such a new identity relates a plural term (variable or constant) to an individual term, and this is quite unorthodox. It is quite easy to alter the language to handle such an hybrid and broader identity. But there is a cost. As a matter of fact identity is normally characterized by Leibniz's law. However, the new CAI identity relation

[50] See Linnebo (2017, §4).

immediately gets in trouble with it, specifically with the law of the indiscernibility of identical. For if the composite objects is identical to its parts, then they must share all and only the same properties. But the composite object is one, whereas its parts are many, and this flies clearly in the face of the indiscernibility of identical. Thus it seems that CAI asks for a very substantial as well as counter-intuitive change of the principles governing identity,[51] and most philosophers are not willing to make it. Secondly, if we accept CAI the proliferation of entities is actually blocked. For the whole composite object is identical to its parts and we don't have to count it as something more than them. Thus, given a collection of n objects, CAI enable us to count exactly n entities – and not $2^n + 1$ – even when Unrestricted Composition holds. Then, CAI makes mereology ontologically innocent, meaning that the mereological axioms we are willing to accept and the theorems we get from them do not affect ontology – i.e. which and how many objects inhabit the domain. Thirdly, it has been proved that the assumption of CAI in a plural logical context makes the relation «is one of» to collapse on parthood, which brings some major problems. To see that, first we need to accept a very reasonable and uncontentious principle called Plural Covering:

$$x < y \to \exists zz\,(y = \bigoplus zz \land x \prec zz) \qquad \text{(Plural Covering)}$$

This principle says that if x is a part of a composite object, y, then it is also one of a plurality of things the sum of which is y. With that, it is possible to prove Collapse, that is:

$$y = \bigoplus xx \to \forall z\,(z \prec xx \leftrightarrow z < y) \qquad \text{(Collapse)}$$

The problem with Collapse is the following. Take a plurality of objects xx such that their fusion is y. According to Collapse, every part of y is one of the xx. But this is plausibly false since you can legitimately have a part of y that is not one of the xx. For instance, my body is the mereological sum of the plurality of the atoms it is made of. But my nose is not one of these atoms, despite it being a part of my body.

In §3.5 we will see how these issues relate also to gluon theory.

[51] A different strategy to solve this problem is to act on the notion of 'property' and show that the numerical properties (and possibly others) are *sui generis* and they are not allowed to be used in the indiscernibility of identical.

3.4 Gluon theory

To solve the one and the many problem Priest (2014b) develops gluon theory (GT). He abandons two key assumptions of the debate, namely that any solution has to be consistent and that identity is transitive. The first assumption is abandoned by the introduction of gluons, which are inconsistent objects. This in turn requires the endorsement of dialetheism – and therefore of paraconsistency – and forces Priest to accept a non-transitive notion of identity. Besides this, Priest also uses a NWFM. However, as I will show, GT implies some undesired consequences which undermine its rational acceptability. Then, let's now construct GT from the bottom and examine its principal details. To do that, we just need to make the following six steps.

Step one: we accept and implement noneism, for which I refer to §3.1.1. Recall that for Priest to be an object is to be identical to something. Hence being a possible object of reference or having at least one property are sufficient for being an object. Given noneism, the domain of objects includes both existent and non-existent objects. We can quantify over objects by using the particular and the universal quantifiers, $\mathfrak{S}$ and $\mathfrak{A}$ respectively, without any existential commitment. We have an existence predicate, E, to signify that the object exists – which, for the record, Priest (2014b, p. xxii) takes to be equivalent with having the potentiality to enter into causal relations. Moreover, we use the same symbols for second-order universal and particular quantifiers, which range over properties. Thus $\mathfrak{S}X\,Xa$ and $\mathfrak{A}Y\,Yb$ are read as «Some property X is had by object a» and «All properties Y are had by object b», respectively.

Step two: we accept and implement dialetheism, for which I refer to §1. To do that, we just need (i) to start from the possible and impossible worlds modal framework we have discussed in §3.1.1, (ii) to impose that the logic of @ (and the other possible worlds) is paraconsistent and (iii) to accept some dialetheias. We may want to accept some semantic or set-theoretic paradoxical conclusion as true, or maybe we could be inclined to accept some of the other Priest's examples of dialetheias we have met in §1.3; but this is not necessary. For we are obliged to accept dialetheism because of the dialetheias concerning some inconsistent actual objects – i.e. gluons – we are going to define in step six. Moreover, we will allow for second order quantification – i.e. we are going to use second-order LP. For a very short introduction to it, see Priest (2014b, §2.10). In addition, as we mentioned in §3.1.1, it is important to notice that Priest has available also a relevant implication, for instance that of BX or some stronger relevant logic. This will prove crucial to define gluons and later to deduce some undesired consequences of GT.

Step three: we define inconsistent objects. An object x is inconsistent (Ix) iff it has at least one contradictory property, Y, i.e. iff x is both in the extension and in the anti-extension of at least one predicate, Y. Formally:

$$\text{Inconsistent object (IO):} \qquad Ix \stackrel{\text{def}}{=\!=} \mathfrak{S}Y(Yx \wedge \neg Yx)$$

As we will show, an inconsistent object both is and is not an object because it is identical to something while it is also not identical to anything.

Step four: we set up a very peculiar notion of identity. Priest uses the standard Leibnizian definition of the identity relation:

$$\text{Identity (ID):} \qquad a = b \text{ iff } \mathfrak{A}X\,(Xa \leftrightarrow Xb)$$

But the material biconditional $Xa \leftrightarrow Xb$ that is used in this definition is understood as defined as $(\neg Xa \vee Xb) \wedge (\neg Xb \vee Xa)$, and since the underlying logic is *LP*, it turns out to be reflexive, symmetric and non-transitive – as well as non-detachable. Thus, identity inherits all these properties from the biconditional. To get a feeling for *LP* and to show that transitivity really does not follow, consider the property of being identical with c, i.e. $x = c$. Here is an example showing that although $a = b$ and $b = c$, we do not have $a = c$. The crucial assumption is that we also have $b \neq c$ (so b is an inconsistent object). Applying the definition of identity to a and b, we have $(a \neq c \vee b = c) \wedge (b \neq c \vee a = c)$. Since the first disjunct of the second conjunct is true, the second conjunct is true. This means we do not have to conclude that $a = c$ in order for the second conjunct to be true. Hence there is an interpretation according to which $a \neq c$ even though $a = b$ and $b = c$ are both true. (Of course, if a, b and c are all consistent objects, then identity *is* transitive.)

Step five: we embrace a specific (even though not complete) account of properties. Under an abundant conception of property, any condition containing a single variable expresses a property. However, for Priest this conception is problematic since he thinks that there are many examples of such conditions that do not specify a property. For example, «being red and Paris is in France» or «$x = x$ or Caesar was a frog» (Priest, 2014b, p. 24). Therefore, Priest operates with a sparser notion of property, although he admits being unable to give necessary and sufficient conditions for determining whether a predicate expresses a property or not. But he does give some useful constraints. A condition with a free variable does not determine a property if its truth conditions at an index of evaluation (world, time, etc.) make reference to another index of evaluation. (This means, roughly, that open sentences containing intensional operators do not specify a property.) The reason is that in intensional contexts truth is not preserved by the material conditional – not even in consistent cases. For example, «Giorgio believes that x is happy» does not express a property because Giorgio can believe this of Clark Kent without believing it of Superman. Other conditions may or may not express a property. So much for properties. [52]

Step six: we define gluons. According to Priest, what solves the one and the many problem and constitutes the unity of an object is its gluon. This object is conceived to be identical to all and only the parts of the unified object. Therefore, the metaphysical glue – i.e. the explanation we were looking for – provided by gluons is their being identical with the parts. A gluon glues the parts of a whole because it is identical to all of them: that is, there is no metaphysical space between a gluon and the parts it glues. Thus, there is no need to ask how it is possible that the gluon and the parts are joined together, because the gluon *is* the parts. Hence, Bradley's regress is stopped.

As in §3.2, let $<$ express the parthood relation. For now we follow Priest (2014b, p. 20, fn. 7) in staying neutral on whether this is parthood or proper parthood. Given any composite object, u, Priest defines its gluon, g_u, as an object which is identical to all and only the parts of u. Formally:

$$\text{Gluon (G):} \qquad y = g_u \overset{\text{def}}{=\!=} \mathfrak{A}x(x < u \leftrightarrow y = x)^{53}$$

And since identity is not transitive in GT, the non-identity of the parts – i.e. their being numerically different – is preserved.

This tells us about the gluon of a composite object. What about the gluon of a simple (i.e. atomic) object? In that case the gluon is the simple object itself. We may call it an improper gluon to distinguish it from the proper gluon of a composite object. And what about the unity of gluons themselves? Priest takes gluons to be simple, i.e. non-partite objects. Thus, the gluon is its own gluon and therefore it has no proper part.

Let us now note some important facts about gluons, which also show that they are contradictory entities.

(F1) Every gluon is self-identical.

[52] In this context, it is interesting to report an appealing remark made by one of the referees. Priest argues against the distinction between nuclear and extranuclear properties, based on the claim that there is no principled, non-ad-hoc way of making such distinction. By analogy, should not this hold also for the distinction between conditions that pick properties out and ones that do not? If yes, there would not be any sparser notion of property, *contra* Priest's assumption – i.e. step five. Also, in that case: could it be possible to solve the problems related to an abundant account of properties by resorting to a liberalized CP for properties and impossible worlds as Priest does with the CP for objects? These sound like very interesting questions that deserve to be explored.

[53] I take $A \leftrightarrow B$ as defined by $(A \rightarrow B) \wedge (B \rightarrow A)$ and I assume that $\rightarrow$ is the relevant implication of BX or some stronger relevant logic, in line with Priest (2014a, fn. 16).

Proof: it follows immediately from the reflexivity of identity.

(F2) Every object has its gluon as a part.

Proof: it follows from (F1) and (G).

(F3) For every object, its gluon is unique.

Proof: suppose that g_u and g'_u are gluons of the same object. Then, we get $g_u < u$ and $g'_u < u$ from (F2). Therefore, from (G) it follows that $g_u = g'_u$.

(F4) Every gluon of a partite object is not self-identical.

Proof (from Zolghadr, 2019, p. 72): let u be a partite object. Then, it has at least two distinct proper parts, p_1 and p_2. Since $p_1 \neq p_2$, there must be at least one property X such that $Xp_1 \wedge \neg Xp_2$ (or *vice versa*). From (G) we get $g_u = p_1$ and $g_u = p_2$. Then, from (ID) it follows that $\mathfrak{S}X(Xg_u \wedge \neg Xg_u)$, therefore $\neg\mathfrak{A}X(Xg_u \leftrightarrow Xg_u)$, therefore $g_u \neq g_u$.

(F5) Every gluon of a partite object is inconsistent.

Proof: consider the property expressed by «to be identical to something», which according to Priest expresses the property of being an object. Then, as Priest (2014b, pp. 20-21) notes, from (F1), (F4) and (IO) it follows that g_u is inconsistent.

(F6) For every partite object, the gluon of the object is numerically distinct from each of the object's parts.

Proof (from Priest, 2014b, p. 22): let u be a partite object with only two proper parts, p_1 and p_2 (the proof immediately extends to objects with an arbitrary number of proper parts). As we know, since $p_1 \neq p_2$ there must be at least one property X such that $Xp_1 \wedge \neg Xp_2$ (or *vice versa*). Thus, $Xg_u \wedge \neg Xg_u$ and hence $\neg(Xp_1 \leftrightarrow Xg_u)$ and $\neg(Xp_2 \leftrightarrow Xg_u)$. Therefore, $p_1 \neq g_u$ and $p_2 \neq g_u$.

(F7) The gluon of a partite object is and is not an object.

Proof: follows immediately from (F1) and (F4) given that Priest takes the property of being an object to be equivalent to the property of being self-identical .

(F8) For every partite object, the gluon of the object is not part of the object.

Proof: it follows from (G) and either (F4) or (F6) using *modus tollens*.[54]

Most of these facts are proven by Priest, except for (F8) which is novel. Note also that (F7) iff (F5). Consider left-to-right first. Assume (F7), then given the definition of being an object we have $\mathfrak{S}x(x = g_u) \wedge \mathfrak{A}x(x \neq g_u)$; hence (F5) follows if you take the property expressed by «is identical to something». The other direction is equal to the proof of (F7). So talking about inconsistent objects is the same as talking about objects that are also non-objects, which in turn is the same as talking about objects that are not self-identical.

Besides the distinction between proper and improper gluon, Priest also distinguishes between prime and non-prime gluons. A prime gluon is a gluon which has all the properties of every part of the object that it unites. A non-prime gluon is a gluon that is not prime. Basically, a gluon is non-prime if at least one of the consistent proper parts of the whole object whose unity it accounts for is also a proper part of another distinct object which partially overlaps with the first one. For consider the following situation described by Priest (2014b, p. 22). Take an object u with two consistent proper parts, a and b, such that b is also a proper part of another object, v, which has also c as a supplementary proper part. Then, u and v are two distinct partially overlapping objects. Also, we have the gluon of u, g_u, and that of v, g_v. From (G) and the consistency of b it follows that $g_u = g_v$, and therefore $g_u < v$ and $g_v < u$. But since the parts of u and v are not all the same, g_u and g_v cannot be prime gluons. If they were, we would get $u = v$, which is not the case.

Before moving further, a short comment on the nature of gluons is in order. For we are left with few clues about what kind of objects gluons are. Of course, they are inconsistent objects. But are they concrete? Abstract? Existent? Nonexistent? The answer is to be found in (G) and (ID), and depends on the nature of the object unified by the gluon under examination. Take c to be a composite concrete object. Then, c has some concrete proper parts and its gluon, g_c, is identical with all and only them. Thus, because of (ID) g_c is concrete. Similarly, take a to be a composite abstract object. Then, a has some abstract proper parts and its gluon, g_a, is identical with all and only them. Thus, because of (ID) g_a is abstract. And the very same result follows in case

[54]Note that (F8) crucially depends on what kind of conditional is meant in (G). As we said in fn. 53 we take → to be the conditional of BX or some stronger relevant logic, so that it is detachable – it validates *modus ponens*, but also *modus tollens*. But had we make a different choice for →, arguably we would have had a different situation. For instance, (F8) does not hold if → is the usual material conditional. For in that case *modus tollens* is LP-invalid.

of abstract or concrete atoms. Moreover, the same line of reasoning can be applied with respect to existent and nonexistent objects. Therefore, the gluon of an existent object is existent, whereas that of a nonexistent object is nonexistent.[55]

So much for gluons. But note that by introducing these unconventional objects we have brought mereology into play. Moreover, further mereological notions – e.g. fusion – will prove essential to define some other very peculiar objects of GT – namely, the objects **everything** and **nothing**. Because of that, we now move to explore Priest's mereology for GT.

3.5 Mereology for gluon theory

I have already disclosed that Priest's mereology for GT is non-well-founded. He gives a reason for going in the direction of a NWFM, but before we discuss it I will formally present his mereology and examine its own adequacy independently from any motivation in support of it.

Priest's axiomatization takes proper parthood ($\ll$) as primitive and defines (improper) parthood and overlap thus:[56]

$$\text{Parthood relation:} \quad x < y \stackrel{\text{def}}{=\!=} x \ll y \vee x = y$$

$$\text{Overlap:} \quad x \circ y \stackrel{\text{def}}{=\!=} \mathfrak{S}z\,(z < x \wedge z < y)$$

He uses overlap and the epsilon operator, εy (to be read as «an object y such that») to define the notion of mereological sum and the general sum operator:

$$\text{Sum:} \quad Sz\Phi x \stackrel{\text{def}}{=\!=} \mathfrak{A}w\,(z \circ w \leftrightarrow \mathfrak{S}x\,(\Phi x \wedge w \circ x))$$

$$\text{General Sum operator:} \quad \bigoplus x\,\Phi x \stackrel{\text{def}}{=\!=} \varepsilon z\, Sz\Phi x$$ [57]

[55] Note that we may have some unusual scenarios in which a gluon turns out to be both abstract and concrete, or both existent and nonexistent. As an example, let assume there is a composite concrete object c which has also one abstract proper part – together with the other concrete proper parts. Thus, its gluon turns out to be both concrete and abstract. Therefore, a consequence we can draw from these considerations is that if we want to prevent gluons from being both abstract and concrete, or both existent and nonexistent, we need to impose the following constraint: (i) composite *abstracta* can only have abstract proper parts, (ii) composite *concreta* can only have concrete proper parts, (iii) composite existent objects can only have existent proper parts, and (iv) composite nonexistent objects can only have nonexistent proper parts. In other words, we want the notions of abstractness and concreteness, as well as those of existence and nonexistence, to be exclusive. However, given the mereology Priest opts for – namely, the principle of unrestricted Composition – we cannot get rid of such unusual situations where we have objects both existent and nonexistent, or both concrete and abstract. More on that in §3.5.

[56] I won't follow Priest (2014b) in using a notation which includes also some set-theoretic notions. I prefer to stay true to the notation I have been using since the beginning of the present chapter. Consequently, my formalization of his mereological principles and definitions will (slightly) differ from his own.

[57] It is worth defining fusion also by using plural variables, since later we will refer to CAI. Then:

Priest takes proper parthood to be transitive and the mereological sum operation to be unrestricted. The caveat here is that just because any two or more objects have a sum does not mean that this sum *exists*, for being a noneist Priest takes existence to be different from being an object. Then, we have the following axioms:

$$x \ll y \land y \ll z \rightarrow x \ll z \qquad (\ll\text{-Transitivity})$$

$$\mathfrak{S}x\,\Phi x \rightarrow \mathfrak{S}z\,Sz\Phi x \qquad \text{(Unrestricted Composition)}$$

Thus, so far Priest's mereological theory is basically the same as Cotnoir and Bacon (2012)'s NWFM – except for the existentially non-committing quantifiers (i.e. the noneistic setting). But now he assumes a principle which makes his mereology different – and inconsistent, as we will see in a moment. He accepts the Extensionality of Overlap, according to which overlapping all the same objects is sufficient for identity:

$$\mathfrak{A}z(z \mathbin{\bigcirc\!\!\!\!\bigcirc} x \leftrightarrow z \mathbin{\bigcirc\!\!\!\!\bigcirc} y) \rightarrow x = y \qquad \text{(Extensionality of Overlap)}$$

And from this we can immediately conclude that given any collection of objects, their fusion is unique – i.e. the principle of Uniqueness of Composition.[58]

However, as it should be clear from what we said in §3.2.1, this kind of mereology is problematic. But before pointing the finger at it, let me say something about one virtue of Priest's theory. In the previous section we saw that the gluon of an object is identical with all the parts of the object. This is a kind of one-many identity that might remind some readers of the Composition as Identity debate. As we saw in §3.3, a central result in that debate is that a strong version of CAI, according to which a whole is literally identical to all its parts (taken collectively), results in a collapse of the parthood relation onto the is-one-of relation ($\prec$) of plural logic. This has various consequences that many[59] are unwilling to swallow. Fortunately for Priest, his mereology does not result in Collapse, as I will now show.

One derivation from CAI to Collapse uses the principle of Plural Covering, which I rewrite here for convenience:

$$z < x \rightarrow \mathfrak{S}yy(x = \bigoplus yy \land z \prec yy) \qquad \text{(Plural Covering)}$$

$$\bigoplus xx \stackrel{\text{def}}{=\!=} \varepsilon z \mathfrak{A}w(z \mathbin{\bigcirc\!\!\!\!\bigcirc} w \leftrightarrow \mathfrak{S}x(x \prec xx \leftrightarrow w \mathbin{\bigcirc\!\!\!\!\bigcirc} x))$$

[58]See Priest (2014b, p. 90).

[59]For example, Sider (2014), Calosi (2016) and Loss (2018).

The argument goes as follows. Suppose x is the fusion of the yy. Then, from right-to-left is easy: let z be one of yy, then by the definition of sum, z is part of x. For the left-to-right direction, suppose z is part of x. Then by Plural Covering, there are some ww such that x is the sum of ww and z is one of ww. By Composition as Identity, $x = yy$ and $x = ww$. Thus by the transitivity of identity, $yy = ww$. Hence, since z is one of ww, z is one of yy. But in Priest's account the left-to-right direction does not go through for two reasons. On the one hand, GT does not entail that x, or the gluon of x, is one-many identical to all the parts of x taken collectively. Instead, the gluon is identical to each part. On the other hand, even if it were to entail that, we would still need the transitivity of plural identity to conclude that $yy = ww$. But if singular identity is not transitive, there is no reason why plural identity would be transitive.[60]

There is nonetheless a problem with Priest's mereology. As Cotnoir (2018) explains (and as we have discussed in §3.2.1), the rejection of Antisimmetry for $<$ (Asymmetry for $\ll$) – that is, the presence of loops – implies giving up Extensionality (of Overlap) or Transitivity (because the latter two entail Antisimmetry: if x and y are parts of each other, then they overlap the same objects, hence, by Extensionality, they are identical). Thus, Priest's mereology does not work.[61]

The solution Cotnoir (2018) suggests is to keep $\ll$-Transitivity and replace Extensionality with Strong Supplementation, so to obtain the very same mereology we examined in §3.2.1: Cotnoir and Bacon (2012)'s NWFM. Thus, in this case Priest has also to leave behind Uniqueness of Composition. Otherwise, instead of giving up Extensionality he could give up Transitivity. But I think this would make mereology redundant given that he accepts a paraconsistent non-well-founded set theory according to which sets are extensional and set formation is unrestricted.[62] So if we take parthood to be non-transitive while leave the fusion operation as both unrestricted and extensional, then $<$ will exhibit the exact same formal features as the membership relation in Priest's set theory. So calling an object a set or a fusion, or saying that it has a part or an element, will then be merely a verbal difference. Furthermore, in §3.7 we will see that there is another good reason for Priest to follow Cotnoir's advice which has everything to do with (the object) **nothing**.

Before concluding this section, there are still two issues I want to discuss: (i) the reason why, according to Priest, GT requires a NWFM, and (ii) the features of two very peculiar objects of (or definable in) GT: the objects **everything** and **nothing**. So let us start with (i). Priest says that we need to go for non-well-foundedness with the mereology of GT because prime gluons can be proper parts of themselves –

[60] A reason I bring this up is that the alternative gluon theory conceived by Yagisawa (2017) does fall prey to Collapse, because in this case the gluon of x is one-many identical to all the proper parts of x. (See also Priest (2017) for a more elaborate comparison of the two gluon theories.)

[61] To be precise, Priest's mereology turns out to be inconsistent. As a dialetheist, he might get along with it. But I don't think this is the kind of contradiction he would be willing to accept.

[62] Priest (2006b, §3).

therefore, asking for parthood loops. To see that, he considers the following situation (see Priest, 2014b, p. 89).[63] Take a prime gluon, g_a, which is the gluon of a composite object, a. Take b to be one of the proper parts of a, and fuse it with g_a. Such a fusion, call it c, is guaranteed by Unrestricted Composition, and it counts as part of a, since both g_a and b are parts of it. And given that $c = b + g_a$ we can conclude that $g_a \ll c$. But seeing as g_a is prime, it has every property that every part of a has. Thus, since c has the property *having* g_a *as a proper part*, then g_a also has the very same property. That is, $g_a \ll g_a$. This means that proper parthood is not asymmetric – and improper parthood is not antisymmetric. So Priest's mereology is not well-founded.

But in my view, the situation is confusing for two reasons. First: the previous argument in defense of parthood loops is mistaken. For given that $c = b + g_a$ we cannot conclude that $g_a \ll c$, but only that $g_a < c$. Or at least, and more precisely, $g_a < c$ is not a conclusion we can draw according to how Priest defines Sum and the other mereological notions involved in this argument. For instance, for every x, he has Sxx and $x < x$, but it is not always the case that $x \ll x$. It can be that $x \ll x$ for some x if we go for a NWFM. But this is exactly what is at stake. Second: the kind of parthood relation (whether proper or improper) involved in (G) is relevant to this matter, and Priest has to take a stand. Should we go for ($G_<$) or ($G_\ll$)?

Gluon ($G_<$): $y = g_u \overset{\text{def}}{=\!=} \mathfrak{A}x(x < u \leftrightarrow y = x)$

Gluon ($G_\ll$): $y = g_u \overset{\text{def}}{=\!=} \mathfrak{A}x(x \ll u \leftrightarrow y = x)$

The main difference we get by choosing between these two options is that ($G_<$) implies $g_u = u$, whereas ($G_\ll$) does not. But note that we have also other consequences. For from (F1) and ($G_<$) we get (F2), but from (F1) and ($G_\ll$) we get ($F2_\ll$):

($F2_\ll$) Every object has its gluon as a proper part.

Now, consider the same case of before (Priest, 2014b, p. 89) and assume ($G_\ll$). Thus, ($F2_\ll$). Therefore, $g_a \ll a$ and then $c \ll a$. Thus, $g_a \ll g_a$. Then, it does seem that ($G_\ll$) – but not ($G_<$) – gives Priest the motivation he wants to go for a NWFM. But ($G_\ll$) seems to clash with the case of atoms. Recall that, according to Priest, any atomic object is its own gluon. But atoms do not have proper parts by definition, and then ($G_\ll$) turns out to be a good definition of gluon only in case of composite objects. And this lack of uniformity would strike me as a point against GT.

[63]Priest's description is not always clear. He speaks of Antysymmetry of proper parthood, whereas it is Asymmetry he has in mind. Moreover, there is a typo in the text: «For example, x might be the mereological sum of g and some other independent part of x» (Priest, 2014b, p. 89). The latter x should be a y. This is just to say that what follows is my personal reconstruction of the case he presents.

However, there is a simple argument in support of a NWFM for GT. We just need a composite object, a, with a prime gluon, g_a, a proper part of a, b, and a proper part of b, c. Thus, given $\ll$-Transitivity $c \ll a$, and since g_a is prime we get $g_a \ll g_a$. Then, what we need is at least one object in the domain which is composed by some *intermediate* proper parts, i.e. proper parts having further proper parts. In other words, we need that not everything is immediately and uniquely composed by atoms, that is a highly plausible scenario.

Finally, we focus on (ii), that is on the objects **everything** (**e**) and **nothing** (**n**). We can think of **e** as «the totality of every object» (Priest, 2014b, p. 54), i.e. the object that every object is a part of. Note that this is essentially the object **u** we have defined in §§3.2 and 3.2.1. But since we are considering it in a different context – namely, the dialetheic and noneistic setting of GT – such an object turns out to have different and unusual features than those it has in classical mereology. Thus, I will keep on using **e** as the name for this universal object of GT, in continuity with Priest (2014b). Let us now take a look at some of its important properties. As Priest notes, **e** is an inconsistent object. For since every gluon of a partite object is an object, it is part of **e**; but since every such gluon is not an object, it is not part of **e**. Then, **e** both has and does not have every such gluon as a part. Therefore, **e** is inconsistent. Moreover, given Priest's mereological perspective, **e** is the fusion of every object. Given noneism, he embraces a version of Unrestricted Composition according to which any collection of objects has a fusion. Bear in mind, though, that such a fusion need not exist, since existence is a predicate of only some objects. Thus, a good definition for **e** is:

everything: $$\mathbf{e} \stackrel{\text{def}}{=} \bigoplus z\,(z < z)^{64}$$

Nothing is «the absence of every thing» (Priest, 2014b, p. 55). As for **e**, **n** is an object since we can refer to it.[65] Besides, **n** is an inconsistent object. For since it is an object it is something, but it is the absence of all things too. So **n** is nothing, i.e. **n** is not something. Therefore, **n** is inconsistent. Priest (2014a, p. 156) defines **n** as :

nothing: $$\mathbf{n} \stackrel{\text{def}}{=} \bigoplus z\,(z \neq z)^{66}$$

[64]Note that an equivalent definition of **e** can be given just in terms of identity by replacing $(z < z)$ with $(z = z)$, since they are both universal conditions. For the latter is included into the former.

[65]**Nothing**'s objecthood is controversial. «Nothing» is often not regarded as a (referring) term but rather as as a quantifier. For discussion of Priest's argument that «nothing» is a term, see Sgaravatti and Spolaore (2018). Here I simply follow Priest, for sake of the argument, in taking «nothing» to refer to the object **nothing**.

[66]Or equivalently, as the fusion of the empty set, $\bigoplus \varnothing$. It is also worth noting that there are other possible definitions of **nothing**. For example, Casati and Fujikawa (2019) define it as the mereological complement of **everything**, which brings about different features compared to those displayed within Priest's account.

n has no parts for it is nothing. Since **n** has no parts it is simple and it is thus its own gluon. **n** is an improper part of itself, $\mathbf{n} < \mathbf{n}$, but it is also not identical with itself and hence not an improper part of itself, $\mathbf{n} \not< \mathbf{n}$.[67]

This, so, is the nature of these two «metaphysical beasts», **e** and **n**. But how do they relate to the universal and the null objects we have discussed in §3.2.1? On the one hand, it immediately follows that **e** is a top element by its own definition. After all, if Priest's mereology can be assimilated to Cotnoir and Bacon (2012)'s NWFM, we already know that the presence of a universal object is guaranteed. Though, since Uniqueness of Composition fails, **e** is one of possibly infinitely many universal objects. On the other hand, there is arguably no reason why **n** should count as a bottom element for the mereological structure of GT. Recall we have to make clear whether we are considering a null object – i.e. $\varepsilon x \mathfrak{A} y\ (x < y)$ – or a strict null object – i.e. $\varepsilon x \mathfrak{A} y\ (x < y \wedge y \not\ll x)$. Therefore, we need to face the following question: can we prove that $\mathfrak{A} y\ (\mathbf{n} < y)$ and/or that $\mathfrak{A} y\ (\mathbf{n} < y \wedge y \not\ll \mathbf{n})$? Well, I can see no way to infer that $\mathfrak{A} y\ (\mathbf{n} < y)$. **n** is the fusion of every inconsistent object and there is no mereological reason why it should be part of every thing, i.e. a (strict) null object. However, the story about **n** is not over. I will have more to say on that in §3.7.

3.6 Every object is inconsistent

In this section I will show that GT entails that every object is inconsistent. As far as I know, the first to point this out were Casati and Fujikawa (2014) in their review of *One*. I reconstruct their argument here. They mention both **everything** and a version of unrestricted mereological composition. Actually, one can use either to get the conclusion that every object is inconsistent.
Casati and Fujikawa (2014, p. 503) note that «according to gluon theory and the unrestricted mereological sum operation, almost every object is contradictory». I take their argument to be as follows. Consider the object **everything**. By definition:

(CF1) $\mathfrak{A} x (x \ll \mathbf{e} \vee x = \mathbf{e})$

Now, according to Priest (2014b, p.55) **e** is inconsistent:

(CF2) $I\mathbf{e}$

Moreover, **e** is a partite object, so it has a gluon (Priest, 2014b, p. 55), $g_{\mathbf{e}}$, which – as shown in §3.4 – is identical and not identical to each of **e**'s proper parts:

[67] Priest (2014a, p. 154).

(CF3) $\mathfrak{A}x(x \ll \mathbf{e} \rightarrow (x = g_{\mathbf{e}} \wedge x \neq g_{\mathbf{e}}))$

So, x is in the extension and in the anti-extension of the property of being identical with $g_{\mathbf{e}}$. Thus, by the definition of inconsistent object:

(CF4) $\mathfrak{A}x(x \ll \mathbf{e} \rightarrow Ix)$

Hence, from (CF1), (CF2) and (CF4):

(CF5) $\mathfrak{A}xIx$

Thus, every object is inconsistent. However, Casati and Fujikawa write that «*almost* every object is contradictory» (my italics). I am not sure about which objects would lie outside the scope of their argument. But note that if monism is true and there is only one object, then this object is simple (it is its own gluon) and then it does not follow that this lonely object is inconsistent – i.e. (CF2) would be false. However, if **everything** is the only object there is, it follows that either **nothing** does not exist – otherwise there are two objects – or that **everything** and **nothing** are one and the same object. Note that monism and noneism make strange bed partners. According to noneism, there are non-existing objects. According to monism there is only one object. Hence, unless noneism is only vacuously true, this combination of views entails that the only thing there is, is a non-existing object. One way out of this particular problem would be to define monism as the view according to which only one object *exists*, while allowing for various non-existing objects. But in that case, **everything** is again a partite object which has one proper part that exists – the only existing object – and various non-existing proper parts. In that case, Casati and Fujikawa's argument goes through as before.

The argument depends on there being an object that is the mereological sum of every object. But the same conclusion can be derived without using **e**, at least if there are two distinct objects. Take two distinct objects, x and y. By unrestricted composition, they have a fusion, z. Now, z has a gluon, g_z, such that $x = g_z$ and $y = g_z$. But, since the gluon of z is also not an object, $x \neq g_z$ and $y \neq g_z$. Hence, x is and is not identical to g_z and x is thus an inconsistent object. (And similarly for y.) But this still leaves it open whether z is an inconsistent object. Note that the gluon of z both is and is not a part of z. Hence, z both has and does not have g_z as a part, and z is thus also inconsistent. Now, by generalising the argument we get that every object is inconsistent.

I think this result comes as a surprise to many. Priest seems to think that some objects are consistent. For example, he writes that the transitivity of identity holds for consistent objects (see Priest, 2014b, p. 20) and that the «non-identity of gluons should hardly be the case for everything» (see Priest, 2014b, p. 24). However, I do not see an obvious way to block this argument. The only way, it seems, to resist the argument is to object that mereological relations and (non-)identity relations do not express legitimate properties – for whatever reason. Indeed, Priest seems to suggest this at some places – for example, Priest (2014b, p. 24 fn. 15). But such properties are often invoked in *One*. For example, Priest explicitly notes that *being identical with something* is usually ruled out as a property in a Leibnizian definition of identity because of triviality but that he need not rule this out, he explains, because the biconditional is non-detachable in LP (see Priest, 2014b, p. 20 fn. 4). As another example, Priest characterizes «being an object» as *being identical to something*:

> What I take *being* to mean here is *being an object*—that is [...], being identical to something. Something is an object iff it has properties. For if it has properties, it is certainly an object; and if it is an object, it has properties—at least the property of being an object.
>
> Priest (2014b, p. 49)

Similarly, his argument against the asymmetry of proper parthood uses the property of *having a gluon as a proper part* (see Priest, 2014b, p. 89). More generally, Priest (2014b, p. xxii-xxiii) is committed to a characterization principle according to which for any condition Px there is some object at some (possible or impossible) world which satisfies Px. A very sparse notion of property would run counter to this. If mereological predicates fail to express properties, then it is unclear how Priest would characterize, for example, the object **e**, since this object is defined as the object that has every thing as a part. For these reasons I will not try to resist the argument. Instead I want to see where the conclusion leads: what follows from the claim that every object is inconsistent?

3.7 Everything and nothing are mutual parts

The fact that every object is inconsistent has far-reaching consequences for the part-whole structure of objects because, if Extensionality holds, it follows that **everything** is identical to **nothing**, i.e. $\mathbf{e} = \mathbf{n}$. Remember, **e** is defined as the fusion of every object, whereas **n** is the fusion of all non-objects. Now, from §3.6 we know that every object is inconsistent, which, as we know from §3.4, means that every object is both self-identical and not self-identical: $\mathfrak{A}x(x = x \wedge x \neq x)$. Thus, every object is a part of **e** and every object is a part of **n**, i.e. **e** and **n** have the same parts. That is to say that every object that overlaps **e** overlaps with **n** and *vice versa*. So, by the definition of fusion and the fact that fusions are, for Priest, unique, $\mathbf{e} = \mathbf{n}$. Note, however, that it is also still the case that $\mathbf{e} \neq \mathbf{n}$ since $\mathbf{n}$ is simple whereas $\mathbf{e}$ is not simple;[68] $\mathbf{n}$ is simple because it does not have any object as a proper part, while $\mathbf{e}$ is not simple because every other object is a proper part of it.

The way out of this conclusion is to drop Extensionality – something Priest has to do anyway if he wants a NWFM, as explained in §3.5. If we replace Extensionality with Strong Supplementation we do not get that $\mathbf{e} = \mathbf{n}$, but we do get that $\mathbf{e}$ and $\mathbf{n}$ are parts of each other, i.e. $\mathbf{e} < \mathbf{n}$ and $\mathbf{n} < \mathbf{e}$. In that case, it follows that $\mathbf{e} \ll \mathbf{n}$ and $\mathbf{n} \ll \mathbf{e}$ (by the definition of improper parthood and that $\mathbf{e} \neq \mathbf{n}$ and not also $\mathbf{e} = \mathbf{n}$). Therefore, because of $\ll$-Transitivity $\mathbf{e} \ll \mathbf{e}$ and $\mathbf{n} \ll \mathbf{n}$. Maybe this is a more palatable consequence.

Dropping Extensionality comes at a price, though. It means that for any set of objects, Σ, there may be more than one fusion of it. A non-extensional mereology does not provide any guidance on the number of fusions that can be formed from a single set of objects. So, in principle, there are infinitely many numerically distinct fusions of the same set Σ. In particular, there may be infinitely many distinct fusions of the set $\{x : x = x\}$ – infinitely many **everything**s – and infinitely many distinct fusions of the set $\{x : x \neq x\}$ – infinitely many **nothing**s – , and all these fusions would be parts of each other (by Strong Supplementation and the fact that every object is inconsistent).

As a toy example to illustrate the problem of dropping Extensionality, consider a fusion of the objects satisfying the condition $x = \mathbf{e} \vee x = \mathbf{n}$, i.e. a fusion of $\mathbf{e}$ and $\mathbf{n}$. Is this object identical with $\mathbf{e}$, with $\mathbf{n}$, with both, or with neither? The mereology no longer tells us because Extensionality does not hold. Neither can we simply apply Priest's Leibnizian definition of identity because even if we would know the properties of $\mathbf{e}$ and $\mathbf{n}$, we would not yet know which of these properties are (not) had by $\bigoplus x(x = \mathbf{e} \vee x = \mathbf{n})$.

Notice also that if $\mathbf{e}$ and $\mathbf{n}$ are mutual parts then the gluon of $\mathbf{e}$ is identical with the gluon of $\mathbf{n}$. To see this, note that $\mathbf{n}$ is part of $\mathbf{e}$ and hence $g_{\mathbf{e}} = \mathbf{n}$. Since $g_{\mathbf{n}}$ is part of $\mathbf{n}$ and $\mathbf{n}$ is part of $\mathbf{e}$, by the transitivity of parthood, $g_{\mathbf{n}}$ is part of $\mathbf{e}$. And since $g_{\mathbf{e}}$ is identical with all of $\mathbf{e}$'s parts, $g_{\mathbf{e}} = g_{\mathbf{n}}$. Hence, since the gluon of an object is its being (Priest, 2014b, p. 51), the being of **everything** is the same as the being of **nothing**. Moreover, the gluon of **nothing** is also identical with **everything**. The reason is that $\mathbf{e}$ is a proper part of $\mathbf{n}$ and since $g_{\mathbf{n}}$ is identical to each proper part of $\mathbf{n}$, $g_{\mathbf{n}} = \mathbf{e}$. (Similar reasoning shows that the gluon of **everything** is identical with **nothing**.) Finally, we also have $g_{\mathbf{e}} = \mathbf{e}$, because $\mathbf{e}$ is a proper part of $\mathbf{e}$ by the transitivity of proper parthood. Furthermore, it seems that both $g_{\mathbf{e}}$ and $g_{\mathbf{n}}$ are prime. Priest (2014b, p. 55) argues for the primeness of $g_{\mathbf{e}}$. The primeness of $g_{\mathbf{n}}$ follows from the definition of being prime: a gluon of x is prime if it has all the properties that each of the parts of x has. Since $\mathbf{n}$ does not have any parts, its gluon (vacuously) has all the properties of all the parts of $\mathbf{n}$.

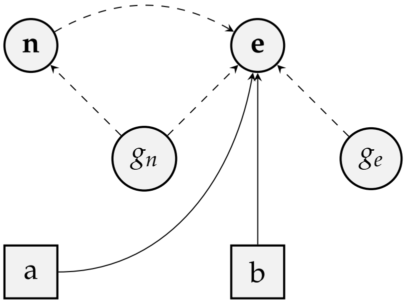
[68] See Zolghadr (2019).

In what follows I present some figures and tables showing in what way proper parthood and identity behave given that every object is inconsistent, and how this is different from what Priest (2014b) seems to suggest. The figures represent the proper parthood relation and the tables represent the identity relation. Figure 3.14a shows how (I think that) Priest takes the proper parthood relation to behave. Square nodes represent consistent objects and round nodes represent inconsistent objects. Arrows represent proper parthood relations and an arrow is dashed if the objects stand both in the extension and anti-extension of the proper parthood relation. Figure 3.14b represent the situation resulting from the fact that every object is inconsistent and assuming that Extensionality fails but Strong Supplementation holds. Table 3.1 shows how Priest takes the identity relation to behave in the case corresponding to the first figure. '$+$' signals that the objects are identical, '$-$' signals that they are not identical, and '$\pm$' signals that they are and are not identical. '$*$' signals that although we know that the objects are in the anti-extension of the identity relation, it is unclear whether the objects also stand in the extension. (They are solely in the anti-extension iff a gluon of x is identical with all and only the *proper* parts of x. But as we know, officially Priest is neutral on the question whether a gluon is identical with the whole of which it is a gluon)[69] Table 3.2 shows how the identity relation behaves in the case represented by the second figure.

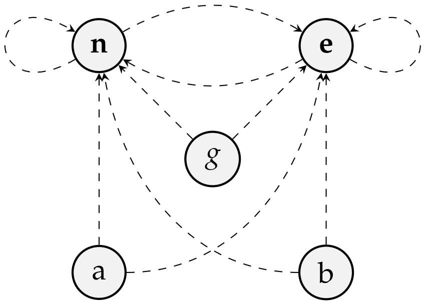
Last, in §3.5 I said that $\mathbf{n}$ is not a (strict) null object. For there is no reason why $\mathfrak{A}y\ (\mathbf{n} < y)$. But from (CF5) – i.e. the inconsistency of every object – we can conclude that $\mathbf{n}$ is a universal object. For $\mathbf{n}$ is the fusion of all inconsistent objects and since $\mathfrak{A}x\ (x \neq x)$ – i.e. (CF5) – $\mathbf{n}$ is the fusion of every object. Thus, every object is part of $\mathbf{n}$. Therefore, $\mathbf{n}$ is a universal object.

[69] See Priest, 2014b, p20, fn.7.

[70] Note that figure (B) shows a partially interpenetrating model, and not a totally interpenetrating model. The former, and not the latter, is what I believe to be Priest's preferred mereological model. For I think he is not willing to accept that, say, his body is part of everything and that everything is part of his body. But some reader of *One* might object that this is in contrast to what he claims in Priest (2014b, §§11-12), where he speaks about *interpenetration*. However, note that in his jargon to say that everything interpenetrate everything does not mean that everything is part of everything, but only that the structural tree of every object is part of the structural tree of every object. In other words, it is the mereological model of the structural trees of the objects to be a totally interpenetrating model, and not the mereological model of the objects.



(A) Priest's preferred model for proper parthood.



(B) Priest's actual model for proper parthood, where **e** and **n** have the same gluon, g.

FIGURE 3.14: Proper parthood relations.[70]

=	**e**	**n**	a	b	g_e	g_n
e	±	−	−	−	−*	−
n	−	±	−	−	−	±
a	−	−	+	−	±	−
b	−	−	−	+	±	−
g_e	−*	−	±	±	±	−
g_n	−	±	−	−	−	±

TABLE 3.1: Priest's preferred model for identity

=	**e**	**n**	a	b	g_e	g_n
e	±	−	−	−	±	±
n	−	±	−	−	±	±
a	−	−	±	−	±	±
b	−	−	−	±	±	±
g_e	±	±	±	±	±	±
g_n	±	±	±	±	±	±

TABLE 3.2: Priest's actual model for identity

3.8 Concluding remarks

To conclude, let me sum up the results of my inquiry. I have shown that there are two ways to make Casati and Fujikawa argument for the inconsistency of every object precise. This means that GT is committed to the claim that every object both is and is not an object. The main consequence of this is that either **everything** is identical with **nothing** or that **everything** and **nothing** are parts of each other. The first disjunct holds if Priest goes for an extensional mereology, whereas the second disjunct holds if he drops Extensionality and instead accepts Strong Supplementation. I think that Priest would prefer the latter, especially because – as Cotnoir (2018) argued – this is the best way to have a NWFM. The only alternative to avoid the conclusion that every object is inconsistent (and the consequences of this claim), seems to be by going for a very sparse notion of property. However, this runs counter to the account of intentionality that Priest favours, which includes a characterization principle that provides an object (in a possible or impossible world) that has the properties that are used to describe it. And regardless of Priest's preferred mereology, I suspect that even those Meinongians who are sympathetic to dialetheism may find it hard to accept that every object is inconsistent or that **everything** and **nothing** are mutual parts (or even identical).

Chapter 4

Towards an inconsistent megethology?

In this final chapter I turn to mathematics. How does dialetheism relates to it? The answer is not straightforward, but the key idea justifying such a kind of inquiry is that there may be mathematical dialetheias. The examples that immediately come to mind are the Russell's and the other set-theoretic paradoxes we have mentioned in §1.3.1. Accepting even one of them as true means revising the cumulative hierarchy given by orthodox set theory to let in also inconsistent sets. Further, besides set theory there may be other branches of mathematics where true contradictions appear, e.g. arithmetic, topology, etc. In other words, the idea worth exploring is that mathematics could be inconsistent – even though not trivial.

As it should now be clear, if we believe in mathematical dialetheias, we must go for a paraconsistent logic to run mathematics. However, not every paraconsistent logic is suitable, and this is for two reasons. On the one hand, the paraconsistent logic we need must be weak enough to avoid triviality caused by the Curry's paradox and related. For naive set and truth theories require more than dropping *ex falso quodlibet* (e.g. dropping also contraction) to be coherent. On the other hand, the paraconsistent logic we need must be strong enough to recapture all the standard (classical) mathematics we know so far. That is, while adding some contradictions to the big book of mathematical truths, we do not want to lose any truth we already jotted down on it.[1]

Many interesting and pioneering works about inconsistent mathematics have been done so far. For example, Brady (1989), Routley (2019) and Priest (2006b, §18) examine and develop an inconsistent (and provably non-trivial) set theory.[2] Arguably, the most complete book collecting the major results in this field is the recently published Weber (2021). But there is still much to be done. As Weber puts it, there is no tome as *Principia Mathematica Paraconsistenta* yet. However, here my aim

[1]Well, that is not mandatory. We might as well give up some theorems of classical mathematics, by virtue of some other benefit we get by opting for a non-classical logic. For example, this is the case of intuitionistic mathematics, where e.g. the Intermediate Value Theorem is not valid, even if it is classically valid. But this move means cutting mathematics as we know it for logical (or philosophical) reasons, and not many – I think – would be willing to do so.

[2]An inconsistent set theory is usually called naive set theory, since it has the naive comprehension schema as axiom.

is not to contribute to such an enterprise. I will not try to recapture a particular piece of mathematics by means of a paraconsistent logic, neither to address some specific mathematical paradox. Instead, I will make a first attempt to outline an inconsistent megethology, that is an alternative foundation for inconsistent mathematics based on (an inconsistent) mereology and plural quantification. Thus, I will quickly discuss what (standard) megethology is first, and then I will try to sketch how an inconsistent megethology could be. But before I start, a preliminary remark is in order: I must confess that I did not succeed in developing such a new foundation. In §4.3 I will present my attempts to get an inconsistent megethology, but unfortunately they turned out to be a wash. I realized that this project is much more challenging than I first thought.[3] However, I decided to show my tries here anyway, together with the reasons why they fail, and some possible course of action to take in order to overcome these problems.

4.1 Megethology

Arguably, set theory[4] is considered the best foundation for mathematics by most. But Lewis (1991) shows that set theory in turn reduces to mereology and the theory of singleton function.[5] Here is the initial part of the introduction, where he summarizes (the first part of) his project:

> There is more mereology in set theory than we usually think. The parts of a class are exactly the subclasses (except that, for this purpose, the null set should not count as a class). The notion of a singleton, or unit set, can serve as the distinctive primitive of set theory. The rest is mereology: a class is the fusion of its singleton subclasses, something is a member of a class iff its singleton is part of that class. If we axiomatize set theory with singleton as primitive (added to an ontologically innocent framework of plural quantification and mereology), our axioms for 'singleton' closely resemble the Peano axioms for 'successor'. From these axioms, we can regain standard iterative set theory.
>
> Lewis (1991, p. vii)

What Lewis means by class, set and null set – and other notions – have to be made explicit, but his construction should now be clear. Furthermore, that's not the whole story! For Lewis (1993) shows that we do not even need to assume the existence of

[3] Basically, this is so because of the weakness of the logic I will be using, i.e. DKQ.

[4] As a matter of fact, there is more then one set theory (e.g. Von Neumann–Bernays–Gödel set theory, ZFC, etc.). Here I refer to the most canonical one, that is ZFC – named after mathematicians Ernst Zermelo and Abraham Fraenkel, and the axiom of Choice.

[5] A singleton is a one-membered class, and a singleton function is a generative function that gives a one-membered class if applied to an object of its domain – individuals and improper classes, in Lewis' jargon (more on this in what follows). However, Lewis (1991) takes singleton as primitive, so that we cannot define it.

at least one singleton function, since this is guaranteed by the assumption of some hypotheses about the size of Reality (i.e. the domain). Thus, set theory – therefore, mathematics – reduces to megethology (mereology and plural quantification) plus such axioms. In what follows, I will go over the main steps of Lewis' reconstruction.

The mereology adopted by Lewis (1991) is exactly the one I presented in §3.2, i.e. **GEM**. Therefore, we have all the pleasant features we have discussed – e.g. extensionality, unrestrictedness and uniqueness of fusions, etc. Now, one of the crucial points is that members are not parts of classes – i.e. membership ($\in$) is not parthood. For it is easy to see that transitivity breaks: we can have $x \in y \wedge y \in z$ but $x \notin z$. Also, a class is not the fusion of its members. Given the ϕs and the ψs, they form the same class iff they are identical. Instead, their fusion can be the same even if they are not identical. Nonetheless, parthood does play a very important role in set theory: if applied to classes, it corresponds to the subset relation, $\subseteq$.

Let us now state some definitions based on the primitive notion of singleton:

Class: any fusion of singletons.

Individual: anything that has no singletons as parts.

Proper class: a class that has no singleton.

Improper class: every class that is not a proper class.

Member: x is a member of y iff y is a class and the singleton of x is part of y.

Null set: the fusion of all the individuals.

Set: something is a set iff either it is a class that has a singleton, or else it is the null set.

Urelement: any individual other than the null set.

Inclusion: y includes x iff (1) x is the null set and y is the null set or a class, or (2) x and y are classes and x is part of y.

Union: The union of one or more things is defined iff each of them is either a class or the null set; it is the null set if each of them is the null set, otherwise it is the fusion of those of them that are classes.

Then, Lewis (1991) accepts the following theses:[6]

First Th.: one class is part of another iff the first is a subclass of the second.

[6]These are not axioms. As explained further below, we will regain such theses precisely from the axioms.

Division Th.: the objects inhabiting the domain are exhaustively divided into individuals and classes.

Priority Th.: no class is part of any individual.

Fusion Th.: any fusion of individuals is itself an individual.

Some facts follow from these theses. For example, the parts of a class are all and only its subclasses,[7] singletons are mereological atoms, and anything that can be a member of a class has a singleton – i.e. the only things that lack singletons are the proper classes.

Next step is to add plural quantification. By referring to pluralities of objects we can define some further important notions:

Infinite: x is infinite iff x is the fusion of some things, each of which is a proper part of another. Otherwise x is finite.

Large/small: x is large iff there are some things such that (1) no two of them overlap, (2) their fusion is the whole of Reality, and (3) each of them contains exactly one atom that is part of x and at most one other atom. Otherwise x is small.

Few/many: suppose we have some things such that some large thing does not overlap any of them. Then they are few iff there is some small thing x, and there are some things, such that (1) x does not overlap the fusion of the former things, (2) each of the latter things is the fusion of one of the former things and one atom of x, (3) for each of the former things, one of the latter things is the fusion of it and one atom of x, and (4) no atom of x is part of two or more of the latter things. Otherwise they are many.

Another way to define infinity is by using Dedekind's strategy: something infinite is something whose atoms correspond 1-to-1 with only some of its atoms.[8] Similarly, we can also redefine «small/large» and «few/many». Something is *small* iff its atoms correspond 1-to-1 with some but not all the atoms in Reality – provided there are many atoms in Reality; otherwise *large*. And some things are *few* iff they correspond 1-to-1 with some but not all the atoms in Reality; otherwise *many*. And some things are *barely many* iff they correspond 1-to-1 with all the atoms in Reality. Also, here are some relevant facts that follow from these definition: (i) mereological atoms are

[7]This is what Lewis (1991, p. 7) calls *Main Thesis*.

[8]Of course, the set-theoretic notion of 1-to-1 correspondence – i.e. bijection – is not yet available. But, as discussed further below, megethology proves capable of expressing it.

finite, (ii) any part of a small thing is small, (iii) any finite thing is small, and (iv) any fusion of a small thing with a finite thing is small.

Now, we set the axioms. There are three axioms concerning the size of Reality, and four axioms regimenting the singleton function:

Hyp. P: If something is small, then its parts are few.

Hyp. U: If some things are small and few, their fusion is small.

Hyp. I: Some fusion of atoms is infinite and yet small.

Functionality: nothing has two different singletons.

Domain: any part of the null set has a singleton; any singleton has a singleton; any small fusion of singletons has a singleton; and nothing else has a singleton.

Distinctness: no two things have overlapping singletons, nor does any part of the null set overlap any singleton.

Induction: if there are some things, if every part of the null set is one of them, if every singleton of one of them is one of them, and if every fusion of some of them is one of them, then everything is one of them.[9]

The axioms for the singleton function – i.e. «x is the singleton of y» – ensure that such a two place predicate is precisely an injective (by Distinctness) function (by Functionality), and that starting from its domain (set by Domain), we get (by Induction) anything there is in Reality, by iterating the operations of singleton and fusion. Also, from Domain and Distinctness it follows that singletons (the images of the singleton function) are mereological atoms. There are other useful facts that can be derived. Here I just list the more relevant ones: (v) a class is a set iff it is small, (vi) every large class is a proper class, and (vii) the things that can be members are exactly the small classes, the null set and the urelements.

The hypotheses P, U and I establish some crucial features about the size of both singular objects and pluralities. They allow to discriminate between infinite sizes, and together impose that the cardinality of Reality is strongly inaccessible:[10]

> So to regain set theory we need to assume Hypotheses U, P, and I. [...] It is easy to see how any two of the constraints can hold. U and P hold, but I fails, if there

[9]This axiom is stated in terms of «the null set». However, as Lewis (1991, p. 96) makes clear, it can be rewritten just in terms of «singleton».

[10]A cardinal number k is *strongly inaccessible* iff (i) it is uncountable, (ii) it is a *strong limit*, i.e. for every cardinal m such that $m < k$, $2^m < k$, and (iii) it is *regular*, i.e. $\mathrm{cf}(k) = k$, where $\mathrm{cf}(k)$ is the smallest cardinal λ such that k is the union of the ordinals smaller than λ.

> are countably many atoms, so that 'small' means 'finite'. U and I hold, but P fails, if there are aleph-one atoms, so that 'small' means 'countable'. P and I hold, but U fails, if there are beth-omega atoms. Making all three hold together is harder. That takes a (strongly) 'inaccessible' infinity of atoms - an infinity that transcends our commonplace alephs and beths in much the same way that any infinity transcends finitude. There will be inaccessibly many atoms, inaccessibly many singletons, and inaccessibly many sets.
>
> Lewis (1993, p. 228)

Then, Lewis (1991, §4) shows that with the definitions and the axioms we have stated, the four theses above and all the axioms of set theory can be proved within this framework (i.e. megethology).[11]

This is a significant result, but Lewis (1991, p. 45) laments our lack of any knowledge about such a mysterious singleton function: «[w]e know nothing [...] about the nature of the primitive relation between things and their singletons». However, this might not be a problem:

> [w]hat we do know, though, is that this relation satisfies certain structural conditions set forth in the axioms of set theory [...]. We needn't pretend to speak unequivocally of the function that takes members to singletons. Rather, any function that conforms to the appropriate conditions shall count as a singleton function. The content of set theory is that there exists some such function.
>
> Lewis (1991, pp. 45-46)

What Lewis (1991) is suggesting here – reluctantly, it should be noted[12] – is the structuralist way through ramseyfication,[13] as it has been done for arithmetic:

> [t]he structuralist about arithmetic needn't scratch his head about the unknown nature of the number-successor relation; the structuralist about set theory needn't scratch his head about the unknown nature of the member-singleton relation. *Any* function that satisfies the stipulated conditions will do.
>
> Lewis (1991, p. 48, emphasis in original)

The reasons why Lewis (1991) is not entirely content even with this approach are two. First, ramseyfication requires quantification over relations. But relations are (usually conceived as) set-theoretical notions, i.e. classes of ordered pairs. Therefore,

[11] See Lewis (1991, pp. 98-107) for the proofs.

[12] Here is a passage showing Lewis' insatisfaction: «Despite all my misgivings over the notion of singleton, I am not fully convinced that structuralist revolution is the right response. I want to carry on examining set theory as we find it. Therefore I leave structuralism as unfinished business» Lewis (1991, p. 54).

[13] About ramseyfication, see for example Lewis (1991, pp. 46-47) and MacBride et al. (2020, §7.2).

we are using a notion that we do not yet have:

> [i]f we need a set-theoretical definition of 'ordered pair', set-theoretical structuralism can succeed only if we understand set-theoretical notions to begin with. Only if we don't need it can we have it.
>
> Lewis (1991, p. 52)

Second, it seems we have to take the existence of at least one singleton function for granted, with no further explanation. Somehow, it remains shrouded in mystery. But Lewis (1993) finds a way out. For the first problem, megethology proves to be capable of simulating quantification over relations, by means of some techniques developed by A. P. Hazen and J. P. Burgess: «[r]oughly speaking, a quantifier over relations is a plural quantifier over things that encode ordered pairs by mereological means» (Lewis, 1993, p. 18). For the second problem, Lewis shows that hypotheses U, P and I guarantee that there is such a singleton function. This is what he calls the Existence Thesis, that is the most important novelty of Lewis (1993) (see Lewis, 1993, pp. 224-225 for its proof). Thus, we have recaptured the whole set theory – through the structuralist approach. This is why Lewis declares that mathematics is megethology.

4.2 Naive set theory

A naive (or inconsistent) set theory – as the one originally conceived by Cantor – is «one which accepts the paradoxes of set theory as part of the theory; [...] it is a theory according to which the Russell class, for example, the class of all those classes which are not self-membered, both does and also does not belong to itself, and thus is perforce an inconsistent theory. But the hope [...] is that it is not a trivial theory on which not just the Russell paradox, but everything, holds» (Routley, 2019, p. 36). The key idea to develop a naive set theory is to buy only two axioms – i.e. Abstraction (or Comprehension) and Extensionality – and then work meticulously on the underlying logic, so as to prevent triviality and at the same time to preserve enough logical power to recapture the standard results of set theory. There is more than one option on the market. I will follow Routley (2019) and Weber (2012)'s choice of using the relevant logic DKQ.

4.2.1 Logic: DKQ

The two triviality-generating principles for the conditional are contraction and weakening. Therefore, our logic must not validate them. To do that, a good strategy is to opt for a weak relevant logic, namely DKQ.[14] For relevant logics invalidates weakening because it is irrelevant. Also, they are naturally paraconsistent, since $\alpha \wedge \neg\alpha \to \beta$ is invalid on relevance grounds, and the basic relevant implication does not contract (as shown in §1.2.2). Below I follow Weber and Cotnoir (2015, Appendix 1) and present the logic as a first-order axiomatic Hilbert system (with redundancies for a better transparency). The following usual definitions are assumed: $\forall$ is $\neg\exists\neg$, $\alpha \vee \beta$ is $\neg(\neg\alpha \wedge \neg\beta)$, and $(\alpha \leftrightarrow \beta)$ is $(\alpha \to \beta) \wedge (\beta \to \alpha)$.

Axioms:

$\alpha \to \alpha$ (identity)
$\alpha \wedge \beta \to \alpha$ (simplification)
$\alpha \wedge \beta \to \beta$ (simplification)
$\alpha \wedge (\beta \vee \gamma) \to (\alpha \wedge \beta) \vee (\alpha \wedge \gamma)$ (distribution)
$(\alpha \to \beta) \wedge (\beta \to \gamma) \to (\alpha \to \gamma)$ (conjunctive syllogism)
$(\alpha \to \beta) \wedge (\alpha \to \gamma) \to (\alpha \to \beta \wedge \gamma)$
$(\alpha \to \beta) \leftrightarrow (\neg\beta \to \neg\alpha)$ (contraposition)
$\alpha \leftrightarrow \neg\neg\alpha$ (double negation elimination)
$\alpha \vee \neg\alpha$ (excluded middle)
$\forall x(\alpha \to \alpha(a/x))$
$\forall x(\alpha \to \beta) \to (\alpha \to \forall x\beta)$ (with no x free in α)
$\forall x(\alpha \vee \beta) \to \alpha \vee \forall x\beta$ (with no x free in α)

Rules:

$\alpha, \beta \vdash \alpha \wedge \beta$ (adjunction)
$\alpha, \alpha \to \beta \vdash \beta$ (modus ponens)
$\alpha \vdash \forall x\alpha$ (universal generalization)
$\alpha \to \beta, \gamma \to \delta \vdash (\beta \to \gamma) \to (\alpha \to \delta)$ (hypothetical syllogism)
$x = y \vdash \alpha(x) \leftrightarrow \alpha(y)$ (substitution)

Structural Rules:

$$\frac{\alpha \vdash \beta}{\alpha \vee \gamma \vdash \beta \vee \gamma} \qquad \frac{\alpha \vdash \beta}{\exists x\alpha \vdash \exists x\beta}$$

This validates argument by cases and structural *reductio*:

[14] DKQ was originally introduced by Routley and Meyer (1976). Actually, they presented a logic called DL. But DKQ is just a quantified version of DL where one of the axioms (D9) is replaced with Excluded Middle. A good and clear presentation is Routley (2019, §§6-7). Note also that if we remove the axiom $(\alpha \to \beta) \wedge (\alpha \to \gamma) \to (\alpha \to \beta \wedge \gamma)$ from DKQ we get the logic Weber (2021) uses for his naive set theory, subDLQ.

$$\frac{\alpha \vdash \gamma \qquad \beta \vdash \gamma}{\alpha \vee \beta \vdash \gamma} \qquad\qquad \frac{\alpha \vdash \neg\alpha}{\vdash \neg\alpha}$$

Let us make some comments. First, $\wedge$ is additive (or extensional) in DKQ – i.e. $\alpha \to \alpha \wedge \alpha$ is valid. Weber (2021, §4.1) shows that this is a problem, since we get triviality thorough *modus ponens* and the Curry sentence. This is essentially the reason why he modifies DKQ to get a more suitable logic called subDLQ. However, this is not a problem for us. For not much change in practice, except for that benefit. Second, as we said, $\to$ does not weaken: $\alpha \nvdash \beta \to \alpha$. But for practical reasons related to some mathematical proofs, it is important to have a conditional that validates weakening. Such a conditional, called enthymematic conditional, $\mapsto$, can be defined in terms of $\to$ and the t constant (the conjunction of all truths, such that $\alpha \dashv\vdash t \to \alpha$):

$$\alpha \mapsto \beta \overset{\text{def}}{=\!=} \alpha \wedge t \to \beta$$

As expected, $\alpha \vdash \beta \mapsto \alpha$, but $\alpha \mapsto \beta \nvdash \neg\beta \mapsto \neg\alpha$.

4.2.2 Naive set theory

To express naive set theory, the language is extended to include a variable binding term forming operator, $\{\cdot : -\}$, and the nonlogical connective $\in$. The two axioms are:

Axiom 1 (Abstraction): $x \in \{z : \phi(z)\} \leftrightarrow \phi(x)$

Axiom 2 (Extensionality): $\forall z(z \in x \leftrightarrow z \in y) \leftrightarrow x = y$

By existential generalization, the abstraction axiom implies the naive comprehension principle:

$$\exists y \forall x(x \in y \leftrightarrow \phi(x))$$

Abstraction and comprehension are naive or unrestricted (i.e. there is no restriction on ϕ), so the set being defined may appear free in its own defining description. Thus, every predicate determines a set.

Given comprehension and extensionality schemes it is easy to develop most of the main features of classical set theory. All the 'basic' standard set theory can be regained (see e.g. Routley, 2019, §8). For example, the axioms of pairing, null set and power set are just instances of the comprehension schema, and their uniqueness is guaranteed by extensionality. But this is just the beginning. For more substantive results concerning 'higher' standard set theory can be obtained. This is not at all straightforward, since some issues about the properties of inconsistent sets have to be settled – e.g. what the cardinality of inconsistent sets is. But once this is done, naive set theory proves capable of entering the transfinite realm. I do not discuss the details here, but for instance Weber (2012) shows that Cantor's theorem, the wellordering problem, the aleph theorem, and the continuum hypothesis are all settled in such a naive set theory, as well as some large cardinal axioms set-theorists have introduced to solve these problems. For an in-depth discussion of naive set theory see Weber (2021, part III, §5).

4.3 Inconsistent megethology

Following Lewis (1993), I will try to develop the framework informally, except for mereology that will be formally introduced. The logic implemented in our framework is the one I introduced in naive set theory: DKQ.

4.3.1 Inconsistent mereology

An inconsistent (or paraconsistent) mereology is a formal theory of parthood and kindred relations developed on a paraconsistent logic. As for naive set theory, what kind of paraconsistent logic we should chose is a sensible matter. However, a good option is to use DKQ, as Weber and Cotnoir (2015, §4) have shown. Here, I present and discuss their inconsistent mereology, that is the one I will be using to construct an inconsistent megethology. Since I already introduced classic mereology in §3.2, I will proceed fairly quickly.

Following Weber and Cotnoir (2015), here are the axioms and the definitions for our paraconsistent mereology:

Axioms:

PM0 The axiomatization of DKQ given in 4.2.1.

PM1 $x < x$

PM2 $(x < y \wedge y < x) \rightarrow x = y$

PM3 $(x < y \wedge y < z) \rightarrow x < z$

PM4 $\exists y\,(y \not< x) \mapsto \exists z\,(z \not\circ x \wedge \forall y\,((y \not\circ x \rightarrow y < z) \wedge (y \not\circ z \rightarrow y < x)))$

PM5 $\exists x\,\alpha \mapsto \exists z\,\mathrm{lub}(z, \alpha)$

PM6 $\exists x \exists y\,(y \neq x) \mapsto \neg \exists x \forall y\,(x < y)$

PM7 $x < y \mapsto \forall z(z \circ x \rightarrow z \circ y)$

Definitions:

Fusion: $\mathrm{lub}(t, \alpha) \stackrel{\text{def}}{=} \forall x\,(\alpha \mapsto x < t) \wedge \forall w(\forall x\,(\alpha \mapsto x < w) \mapsto t < w)$

Proper parthood: $x \ll y \stackrel{\text{def}}{=} x < y \wedge \exists z\,(z \circ y \wedge z \not\circ x)$

Overlap: $x \circ y \stackrel{\text{def}}{=} \exists z\,(z < x \wedge z < y)$

Disjointness: $x \not\circ y \stackrel{\text{def}}{=} \neg(x \circ y)$

Underlap: $x \odot y \stackrel{\text{def}}{=} \exists z\,(x < z \wedge y < z)$

Let us make some comments. First of all, PM1-6 plus the definitions represent the non-standard axiomatization for **GEM** given by Hovda (2009), that proves to be equivalent to the more standard one I introduced in §3.2.[15] Let us interpret PM1-6 classically and see what they say. Classical PM1-3 fix the meaning of $<$ and ensures it is a partial order. By definition, $\ll$ becomes a strict partial order. Classical PM4 states that if y is not part of x, then there is an object z made up of all and only the non-x-overlapping parts of y. We can call z the complement of x, that is $\overline{x}$. Moreover, the antecedent allows there not to be a complement to the whole universe. This would be the null object – i.e. the object that is part of everything – which we do not have in **GEM**. Also, note that fusion and proper parthood are defined differently than in chapter 3. If the underlying logic is classical first-order logic, these different definitions prove to be equivalent to the previous ones. (But soon we will examine how things change with a paraconsistent logic.) Specifically, a fusion of the αs is their *least upper bound* (lub, for short): a lub of the αs is an object that has all the αs as parts, and is part of any other upper bound of the αs. Classical PM4 guarantees the existence of the fusion for any definable collection of αs – i.e. composition is unrestricted – that is provably unique. Finally, classical PM6 explicitly rules out the null object. But of course, we are not classical here. So, how do the logic DKQ alter the mereological structure?

For a start, consider the structure of figure 4.1 which is a model of Weber and Cotnoir (2015)'s inconsistent mereology, i.e. a **PM**-model. Nodes are objects, thick arrows represent proper parthood relations ($\ll$) and dashed arrows correspond to what we can call the non-proper parthood relations ($\not\ll$). PM1-3 are satisfied, since the graph is intended to be closed under reflexivity and transitivity of parthood, and because there are no mutual parts. Every object of the model has a complement, therefore PM4 is true. PM5 is also true, because there is a fusion for any combination

[15] Well, that is not entirely correct. Of course, you have to replace PM0 with an axiomatization of classical first-order logic plus identity. Consequently, you do not have two different conditional, but just the material implication.

of objects. PM6 is satisfied. To see that, consider we have more than two objects, and therefore $\neg\exists x \forall y (x < y)$. But this is equivalent to $\forall x \exists y (x \not< y)$. Since every object is not a proper part of at least one other object, and because they are distinct, PM6 is true. Finally, PM7. This is a theorem of **GEM**, but it cannot be proved from PM0-6, because PM4 is expressed with a conditional ($\mapsto$) weaker than the material one. Then, Weber and Cotnoir (2015) take it as an axiom. It is easy to check that even this last axiom is true in the model. For example, consider x and y to be instantiated by b and a, respectively. Thus, the only $z \circ a$ is b, which of course overlaps itself. And whatever combination of x and y you pick in the model, the axiom is satisfied.

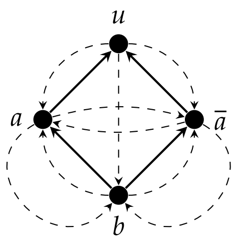


FIGURE 4.1: A simple model of Weber and Cotnoir (2015)'s inconsistent mereology.

Unlike with **GEM**, this model shows that we can have a bottom element (b) in our inconsistent mereology. But this is not the null object in the classical sense. Classically, the null object is part of everything, and there is no object of which it is not a part. Instead, the bottom element in our paraconsistent mereology is part of everything, but it is also not a part of some objects. Furthermore, note that a and $\bar{a}$ are both disjoint and overlapping, that is $a \circ \bar{a} \wedge a \not\circ \bar{a}$. Of course, this is inconsistent! But it is not trivial, since it is not the case that every object both is and is not part of every other object. In other words, the model has an inconsistent but coherent mereological structure.

Let us now note some relevant mereological facts of our inconsistent mereology:

F1$_{PM}$) It is a theorem that $x < y \mapsto x \circ y$.

F2$_{PM}$) Extensionality of proper parthood[16] does not hold. For we can have non identical objects composed by the same proper parts. This is the case of a and $\bar{a}$ in the model of figure 4.1. For being identical, two objects also need to have exactly the same non-proper parts. And, again, this is the case of a and $\bar{a}$. Thus, a and $\bar{a}$ are both identical and not identical.

F3$_{PM}$) Both Weak and Strong Supplementation principles fail: composite objects can have just one proper parts. This is the case of both a and $\bar{a}$ in the model of figure 4.1.

[16] That is, having all and only the same proper parts is both sufficient and necessary for being identical. This is a (controversial) theorem of **GEM**.

F4$_{PM}$) Extensionality of overlap[17] does not hold. For we can have non identical overlapping the same objects. Again, this is the case of a and $\overline{a}$ in the model of figure 4.1.

F5$_{PM}$) The fusion of a collection of objects is unique. For it is a theorem that $\forall u \forall v\, (\text{lub}(u, \alpha) \wedge \text{lub}(v, \alpha) \rightarrow u = v)$.

F6$_{PM}$) The existence and the uniqueness of a *greatest lower bound* (glb) are guaranteed (by PM5 and the uniqueness of fusion), where $\text{glb}(t, \alpha) \stackrel{\text{def}}{=\!=} \forall x\, (\alpha \mapsto t < x) \wedge (\forall w \forall x\, (\alpha \mapsto w < x) \mapsto w < t)$.

F7$_{PM}$) From PM5 follows the existence of a universal object. Given uniqueness of fusion, such a universal object is unique.

F8$_{PM}$) Note that the existence of a bottom element is not guaranteed – i.e. we can have **PM**-models without it. For the complementation axiom PM4 makes sure the existence of the complement of x conditionally, just in case there is something that is not part of x. But it is not **PM**-necessary that there is something that is not part of the universe. Thus, the bottom element is not guaranteed. But if we need it, we can have it by introducing a further axiom, as in Casati and Fujikawa (2019, §5) where they present their own inconsistent mereology: $\mathbf{PM^C}$. This additional axiom is: $\exists x \exists y\, (x \nless y \wedge \text{lub}(y, z = z))$. It ensures that something is not part of the universe (defined as the fusion of all self-identical objects, i.e. all the objects). Therefore, by PM4 the bottom element is guaranteed as the complement of the universal object, $\overline{\text{lub}(t, x = x)}$.

Finally, a very important remark is that such an inconsistent mereology is proved to be nontrivial (see Weber and Cotnoir, 2015, Appendix 2). As we know, this is always a crucial feature for a theory: had been trivial, it would have been useless.

4.3.2 Adding plural quantification

Let us start with plural logic as was introduced in §3.3. Its axioms are Plural Comprehension, Non-Emptiness and Extensionality for Pluralities. Of course, they depend on the logic we adopt. What then if we want to have them in our inconsistent mereology? The only changes we need to consider are where the conditionals occur. Take Plural Comprehension first. Following the naive comprehension schema – that can be derived from Abstraction – I use the stronger contraposible (bi)conditional, $\rightarrow$, in the consequent. Instead, it is better to have the enthymematic conditional as the main connective of the axiom, since it weakens and then allows to add further conjuncts in the antecedent. Second, the Extensionality of Pluralities. For the same reasons,

[17] That is, overlapping all and only the same objects is both sufficient and necessary for being identical. This is a (controversial) theorem of **GEM**.

the two biconditionals will be the contraposible ones, whereas the main connective of the axiom will be the enthymematic conditional. Then, we have:

$$\exists x\, \Phi x \mapsto \exists yy\, \forall x\, (x \prec yy \leftrightarrow \Phi x) \qquad \text{(Plural Comprehension)}$$

$$\forall xx\, \exists y\, (y \prec xx) \qquad \text{(Non-Emptiness)}$$

$$\forall xx\, \forall yy\, (\forall z\, (z \prec xx \leftrightarrow z \prec yy) \mapsto (\Phi xx \leftrightarrow \Phi yy)) \qquad \text{(Extensionality for Pluralities)}$$

4.3.3 Naive set theory regained?

The question we need to answer now is the following: are there any axioms and definitions we can state in our inconsistent megethology from which we can get back the axioms of naive set theory? For since the underlying logic of our inconsistent megethology is precisely that of the naive set theory we have targeted, what we just need to show to regain naive set theory is that its axioms can be derived from those we introduce in our inconsistent megethology.

Let us use the same definitions and the four axioms regimenting the singleton function of Lewis (1991)'s megethology, and see where they lead by using DKQ. Actually, a change is required. In naive set theory we basically have whatever set we can define, since the comprehension schema is unrestricted. That means that we can have really huge sets that do not fit in the standard hierarchy. Thus, to get such objects we need to extend the domain of the singleton function, by allowing the singletons of improper classes. Therefore, I opt for the following axiom that replace Domain:[18]

Domain*: any part of the null set has a singleton; any singleton has a singleton; any fusion of singletons has a singleton. Or, more simply, anything has a singleton.

[18]However, this may not be the only change needed. For example, one referee suggested the following line of reasoning. Due to the paraconsistent setting that allows for non-self-identical objects, Functionality might also break down. Suppose that a and b are two different individual objects, and both are not self-identical (i.e. inconsistent objects): $a \neq b$; $a \neq a$; and $b \neq b$. Intuitively, a is not a member of the fusion of $\{a\}$, since a is not a (but, a is a member of the fusion of $\{a\}$, since a is a). So nothing is a member of the fusion of $\{a\}$ (but, a is). The same would hold for b and the fusion of $\{b\}$. If Extensionality holds, it follows that the fusion of $\{a\}$ is not the fusion of $\{b\}$, but the fusion of $\{a\}$ is the fusion of $\{b\}$ (since nothing is a member of them). If the fusion of $\{a\}$ is $\{a\}$ and the fusion of $\{b\}$ is $\{b\}$, then $\{a\}$ is not $\{b\}$, but $\{a\}$ is $\{b\}$. Now, since $\{a\}$ is $\{b\}$ and $\{a\}$ is the singleton of a, $\{b\}$ is the singleton of a. But, since $\{a\}$ is not $\{b\}$, a has two distinct singletons – Functionality does not hold. Lucky for us, Extensionality does not hold in our inconsistent mereology (recall $\mathbf{F2}_{PM}$), and such an example does not affect our framework. But it is not so obvious that there are no problems with the remaining axioms. Some more careful considerations might be required.

From these axioms, we want to derive those of naive set theory. Let us focus on Extensionality first.

Extensionality: $\forall z(z \in x \leftrightarrow z \in y) \leftrightarrow x = y$

Let us rephrase it as follows: for every z, the singleton of z is a part of x iff the singleton of z is a part of y, iff x is identical with y.

Proof:

Right-to-left. Assume $x = y$. By indiscernibility of identicals, all and only the singletons that are part of x are also part of y. Therefore, $x = y \rightarrow \forall z(z \in x \leftrightarrow z \in y)$ by deduction theorem.

Left-to-right. Assume that for every z the singleton of z is a part of x iff the singleton of z is a part of y. In other words, x and y have exactly the same singletons as parts. By definition of class, x and y are the fusions of their singletons. Since to get x and y we fuse the very same singletons, by uniqueness of fusion $x = y$. Therefore, $\forall z(z \in x \leftrightarrow z \in y) \rightarrow x = y$ by deduction theorem.

But this proof is invalid.[19] For DKQ does not have a general conditional introduction rule, and more generally the deduction theorem is not unrestrictedly valid.[20] This particular feature makes the proofs of conditional formulas very difficult in DKQ, so much so that the same problem will occur even with Abstraction. What to do about that? There are restricted versions of the general conditional introduction rule that Brady (1989) worked out where you have to keep careful track of subscripts. There might be a chance that with Brady's index system the proof works, but it still needs to be verified. If not, a different option could be to strengthen the logic DKQ by adding such a rule of inference. But then, we would need to check if such a new logic is coherent and if it allows to develop non-trivial suitable versions of naive set theory and inconsistent mereology. Unfortunately, this is all I have to offer at the moment.

Let us now move to the second task: deriving Abstraction.

Abstraction: $x \in \{z : \phi(z)\} \leftrightarrow \phi(x)$

[19] Thanks to Aaron Cotnoir and Zach Weber who pointed out the problem. They replied an email by Jeroen Smid – with whom I was collaborating – about another proof we made in DKQ, and showed that it was invalid because of a step where the deduction theorem was involved. For the very same reason, the proof I sketched for Extensionality is not valid.

[20] See Gonzalez and Tapia-Navarro (2021, p. 509).

Let us rephrase Abstraction as follows: given any condition ϕ, the singleton of any object x is part of the class of the ϕs iff x satisfies ϕ.

Proof:

Right-to-left. Assume that $\phi(x)$. Take the pluralities of the ϕs, that is the zz such that $\forall z(z \prec zz \leftrightarrow \phi(z))$. Thus, x is one of the ϕs. By Domain*, all ϕs has a singleton. By unrestricted composition, we have the fusion of all such singletons. By definition of class and uniqueness of fusion, such a fusion is the class $\{z : \phi(z)\}$. By definition of fusion, x's singleton is part of $\{z : \phi(z)\}$. Finally, by deduction theorem we get $\phi(x) \rightarrow x \in \{z : \phi(z)\}$.

Now, to the left-to-right direction. To begin with, let us assume the following plausible LEMMA (I will say more about it in a few lines):

LEMMA: Singletons are atoms.

Left-to-right. Assume x's singleton is part of the class of the ϕs. By definition, the class of the ϕs is the fusion of the singletons of every z such that $\phi(z)$. Then, by assumption x's singleton is part of such fusion. Now, note that singletons are mereological atoms by LEMMA. Thus, x's singleton does not overlap with any other singleton. But here it seems we get stuck. We would normally proceed by *reductio* in the following way. If $\neg\phi(x)$, x is not one of the ϕs. But since x's singleton is a mereological atom, it cannot be part of the fusion of the singletons of the ϕs. If it was, it would overlap at least one of the atoms of that fusion, and therefore it would not be itself an atom. Therefore, $\phi(x)$ by *reductio*. Finally, by deduction theorem we get $x \in \{z : \phi(z)\} \rightarrow \phi(x)$. But these last steps are mistaken. First, we cannot use *reductio*. Second, and likely worse, it is not clear whether from distinct non-overlapping pluralities we get different fusions. For our paraconsistent mereology might not validate the following principle:

$$\forall t\, \forall xx\, \forall yy\, ((\forall z(z \prec xx \vee z \prec yy \rightarrow \text{Atom}(z)) \wedge \text{lub}(t, xx) \wedge \text{lub}(t, yy) \rightarrow xx = yy).$$

In other words, the fusion of the singletons of the ϕs (i.e. the class $\{z : \phi(z)\}$) could have x's singleton as part even if it is an atom distinct from all the singletons of the ϕs.

Since we cannot use *reductio*, we need to get to $\phi(x)$ directly. Before doing that, let me state two principles I am going to use:[21]

[21]Here I use $\circ$ instead of ⚭ because of some LaTeX-related issues. I am sorry for my bad mixed notation.

P$_1$) $x \circ \iota t \,\mathrm{lub}(t, \psi(y)) \mapsto \exists y(x \circ y \wedge \psi(y))$

P$_2$) $\forall x \forall y(\mathrm{Atom}(x) \wedge \mathrm{Atom}(y) \wedge x \circ y \rightarrow x = y)$

P_1 says that whenever an object x overlaps the fusion of some objects satisfying a given condition ψ, then x overlaps at least one of the ψs. P_2 ensures that no two distinct mereological atoms can overlap.
Then, here is a different and more formal try:

Left-to-right (2nd try).

1	(1)	$\{a\} < \{z : \phi(z)\}$	Assumption
1	(2)	$\{a\} < \iota t \,\mathrm{lub}(t, \Phi(y))$	1: class def. + uniqueness of fusion $\Phi(y) := \exists z(y = \{z\} \wedge \phi(z))$
1	(3)	$\{a\} \circ \iota t \,\mathrm{lub}(t, \Phi(y))$	2, $\mathbf{F1}_{PM}$: MPP
1	(4)	$\exists y(\{a\} \circ y \wedge \Phi(y))$	3, P_1: MPP
	(5)	$\forall y(\mathrm{Atom}(\{a\}) \wedge \mathrm{Atom}(y) \wedge \{a\} \circ y \rightarrow \{a\} = y)$	P_2: universal instantiation
1	(6)	$\exists y(\Phi(y) \wedge \{a\} = y)$	4, 5, LEMMA, def. of Φ
1	(7)	$\Phi(\{a\})$	6: indiscernibility of identicals
1	(8)	$\phi(a)$	7: by def. of Φ
	(9)	$\{a\} < \{z : \phi(z)\} \rightarrow \phi(a)$	1, 8: deduction theorem
	(10)	$\{x\} < \{z : \phi(z)\} \rightarrow \phi(x)$	9: universal generalization

Let us make some comments. First, in step (6) I did several steps all at once. To unpack them, consider that from (4) we know that {a} overlaps one of the Φs, y. But from the definition of Φ, we know that such y is a singleton. Therefore, y is an atom by LEMMA. Thus, the antecedent of (5) is true, and by MPP we get the consequent, $\{a\} = y$. Then, since y is one of the Φs, we get (6). Second, this proof crucially depends on LEMMA, P_1 and P_2 being valid, i.e. **PM**-theorems. But is that so?

Unfortunately, the answer is negative at least for P_2. For the model of the figure 4.1 is a counterexample for such a principle: a and $\overline{a}$ are distinct mereological atoms, and yet they are overlapping. Moreover, I currently do not know whether P_1 and the LEMMA are valid. For instance, consider the following partial sketch for proof of the latter:

LEMMA: Singletons are atoms.

(Partial sketch of a) Proof:

By Induction, we know that everything is either a part of the null set, a fusion of singletons, or a fusion of a part of the null set and some singletons. Therefore, if a singleton has parts, they are among these three kinds of objects. However, by Distinctness we know that either no part of the null set or any singleton except itself can be part of a singleton. But here I get stuck. I would normally proceed by *reductio* and say that it cannot even have any fusion as part, since such fusions have some singleton or part of the null set as part, and by transitivity this part would be also part of the singleton, contrary to Distinctness. But we cannot use *reductio* here, and I do not have a direct alternative way of proving that a singleton cannot have any fusion as part.

As if that weren't enough, we still have the very same problem with the deduction theorem we encountered before: it is not generally valid, so that the last step of the right-to-left proof and the step (9) of the left-to-right proof are not allowed. About that, what I said before still applies.

Aside from the problem with the deduction theorem, is there a way out to get a proof of Abstraction? As far as I see, there are two options we can go for. The first one is to find a different proof that does not resort to LEMMA, P_1 and P_2. Unfortunately, I have nothing to offer about that at the moment. The second option[22] is to make some changes to Weber and Cotnoir (2015)'s inconsistent mereology so that LEMMA, P_1 and P_2 become valid. In particular, we could add all of them as axioms and find a non-triviality proof for such a new version of inconsistent mereology. This may be a viable solution worth exploring.

Be that as it may, what I think I have shown in this chapter is that such a project is far from easy, the main problem being the weakness of DKQ. But let me conclude with a final thought. Assume that there is a way to overcome the issue concerning the conditional introduction rule, and that a non-trivial inconsistent mereology validating the proof I sketched before for Abstraction is actually available. In that case, note that I did not need any of the Hypotheses P, U and I. Therefore, the only axioms we need to buy for our inconsistent megethology are those regimenting the singleton function – together with the logical and the mereological ones, and those concerning pluralities. But there might be a better option. Following the second part of Lewis' work – i.e. Lewis (1993) –, it might be possible to find some other different and 'less mysterious' axioms from which to show that the existence of one such singleton function is guaranteed – likewise Lewis' Existence Thesis follows from Hypotheses P, U and I. Besides, Burgess and Hazen's techniques to simulate quantification over

[22]This strategy was suggested by Graham Priest in a conversation we had during my visiting period at CUNY Graduate Center, in March 2022.

relation might work even in our inconsistent megethology. In that case, a structuralist reconstruction of naive set theory would succeed. But that is material for another story.

Conclusion

A rational evaluation of dialetheism by considering all the relevant *criteria* and comparing its overall score with that of its main competitor (i.e. the view that there are not true contradictions) is far from being complete – if it can ever be. What I hope I have shown with my work is that we have some more reasons to resist dialetheism. First, as argued in chapter 2, the exclusion problem still needs a solution, since Priest's pragmatic approach gives rise to revenge. This can be interpreted as a weakening in terms of the explanatory power of dialetheism, since it cannot account for some crucial *phenomena* that seem to require exclusivity. Second, chapter 3 shows that its alleged fruitful application to solve the one and the many problem is not so plainly beneficial. Some undesired consequences can be derived from gluon theory. This does not mean that dialetheism is wrong, but just that it is not so obvious that it helps to provide an effective solution to such a metaphysical problem. Maybe, some mereological adjustments to gluon theory can avoid such bad consequences, but this has to be seen. Thus, this point can be interpreted as a rejection of the explanatory power ascribed to dialetheism – passing through gluon theory – with regard to this very specific metaphysical problem. Finally, chapter 4 does not affect the rational evaluation of dialetheism. Whether or not an inconsistent megethology is possible has not influence over such an evaluation, since it does not add anything to naive set theory. What does positively affect the evaluation, instead, are the results obtained by the first inconsistent mathematicians. They showed that inconsistent mathematics is a coherent and possibly very powerful enterprise, and that it might be reasonable to make room for mathematical dialetheias in the big book of mathematics.

In the last three months of my doctoral research I had the opportunity to visit Prof. Graham Priest at City University of New York. I had the privilege of discussing my research with him.[23] Most of his comments have already been implemented in the previous chapters, but some other important ones have not. Therefore, I would like to conclude my thesis with two quick but relevant thoughts he gave me with respect to chapters 2 and 3. First, he does not accept DLEAC as a paraconsistent logic that properly captures his dialetheic view. After all, we do have an exclusive operator embedded in DLEAC: the star operator (∗). And it does not matter if we interpret ∗ as a pragmatical notion: since we set it inside our logical machinery, it count as a logic operator, not a pragmatic one. For a boolean negation can be defined as $\underbrace{\neg\alpha}_{(*)}$ in DLEAC. This point deeply affect my formal arguments, but not my informal ones.

[23]To be precise, only the first three chapters of my thesis.

Therefore, the informal versions of the denial and the rejection paradoxes still work. But as expected, he takes them to be cases of rational dilemmas. Second, he said Priest (2014b) is silent about whether or not his preferred mereology for gluon theory includes the axiom of unrestricted composition. And he made clear that there is no reason why noneism should force him to buy such axiom. Nonetheless, not much in my analysis depends on it and the consequences I draw are still valid. However, Priest pointed out that he has many moves available to fix these problems. For instance, a change of the kind of (bi)conditional occurring in (G) – the definition of gluon – may have a strong impact. Some of the facts we proved in §3.4 strictly rely on the formal features of implication, and therefore it might be possible to avoid the consequence that every object is inconsistent. Also, giving a precise theory of properhood could help to solve the same problem (as already mentioned in §3.6). And finally, some mereological changes – such as dropping Strong Supplementation, as well as Extensionality, or finding a different definition for **n** – may prove effective to avoid **e** and **n** being mutual proper parts.

Before concluding, let me spell out that in my mind I have not yet settled the issue of the rational acceptability of dialetheism. Even if here I tried to give some reasons to resist it, I still do not know if I can tell myself a dialetheist or a non-dialetheist. And that is precisely because there are some reasons pointing towards the 'dialetheic horizon' that strike me as quite compelling – e.g. the semantic closure, the dialetheias at the limits of thoughts, and others. A complete and agreed evaluation of such a view is likely an hardly achievable task. Maybe, only the course of events will determine its rational success.

Bibliography

Baxter, Donald LM (1988). "Many-one identity". In: *Philosophical Papers* 17.3, pp. 193–216.

Beall, JC (2006). "Review of G. Priest, Towards Non-Being: The Logic and Metaphysics of Intentionality". In: *Notre Dame Philosophical Reviews* 2006.9.

Beall, Jc (2014). "End of inclosure". In: *Mind* 123.491, pp. 829–849.

— (2016). "Do laws deliver gluts?" In: *Law and the New Logics*. Ed. by Glenn and Smith. Cambridge: Cambridge University Press, pp. 199–207.

Beall, JC, Michael Glanzberg, and David Ripley (2018). *Formal theories of truth*. Oxford University Press.

Beall, Jc and Julien Murzi (2013). "Two flavors of Curry's paradox". In: *The Journal of Philosophy* 110.3, pp. 143–165.

Berto, Francesco (2006). "Meaning, metaphysics, and contradiction". In: *American Philosophical Quarterly* 43.4, pp. 283–297.

— (2008). "Modal meinongianism for fictional objects". In: *Metaphysica* 9.2, pp. 205–218.

— (2012). "How to Rule Out Things with Words: Strong Paraconsistency and the Algebra of Exclusion". In: *New waves in philosophical logic*. Ed. by Greg Restall and Gillian Russell. Springer.

— (2014). "Absolute contradiction, dialetheism, and revenge". In:

— (2017). "Impossible worlds and the logic of imagination". In: *Erkenntnis* 82.6, pp. 1277–1297.

Bolander, Thomas (2017). "Self-Reference". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2017. Metaphysics Research Lab, Stanford University.

Boolos, George (1984). "To Be is to be a Value of a Variable (or to be Some Values of Some Variables)". In: *The Journal of Philosophy* 81.8, pp. 430–449.

— (1985). "Nominalist platonism". In: *The Philosophical Review* 94.3, pp. 327–344.

Brady, Ross T. (1989). "The non-triviality of dialectical set theory". In: *Paraconsistent Logic: Essays on the Inconsistent*. Ed. by G. Priest, R. Routley, and J. Norman. Munchen: Philosophia Verlag.

Burgis, Ben and Otávio Bueno (2019). "Liars with Curry: Dialetheism and the Prospects for a Uniform Solution". In: *Dialetheism and its Applications*. Springer, pp. 1–20.

Calosi, Claudio (2016). "Composition as Identity and Mereological Nihilism". In: *The Philosophical Quarterly* 66.263, pp. 219–235.

Cantor, Georg (1882). “Über unendliche, lineare Punktmannichfaltigkeiten”. In: *Mathematische Annalen* 20.1, pp. 113–121.

Carrara, Massimiliano, Filippo Mancini, and Jeroen Smid (2021). “On the Overlap Between Everything and Nothing”. In: *Logic and Logical Philosophy*, pp. 1–18.

Carrara, Massimiliano, Filippo Mancini, and Andrea Strollo (2021). “DLEAC e il paradosso del diniego”. In: *Dialeteismo e storia della filosofia (occidentale)*. Padova University Press.

Carrara, Massimiliano and Enrico Martino (2019). “DLEAC: A Dialetheic Logic with Exclusive Assumptions and Conclusions”. In: *Topoi* 38.2, pp. 379–388.

Casati, Filippo and Naoya Fujikawa (2014). “Review of ‘Graham Priest, *One: Being an Investigation into the Unity of Reality and of its Parts, including the Singular Object which is Nothingness*’”. In: *Logic and Logical Philosophy* 23.4, pp. 499–508.

— (2019). “Nothingness, Meinongianism and inconsistent mereology”. In: *Synthese* 196.9, pp. 3739–3772.

Chisholm, Roderick M (1973). “Parts as essential to their wholes”. In: *The Review of Metaphysics*, pp. 581–603.

Cotnoir, Aaron J (2013). “Beyond atomism”. In: *Thought: A Journal of Philosophy* 2.1, pp. 67–72.

— (2016). “How Many Angels Can Be in the Same Place at the Same Time? A Defence of Mereological Universalism”. In: *Mind* 125.500, pp. 959–965.

Cotnoir, Aaron J. (2018). “A Note on Priest’s Mereology”. In: *The Australasian Journal of Logic* 15.4, pp. 642–645.

Cotnoir, Aaron J and Andrew Bacon (2012). “Non-wellfounded mereology”. In: *The Review of Symbolic Logic* 5.2, pp. 187–204.

Cotnoir, AJ and Achille C Varzi (2021). *Mereology*. Oxford University Press.

Crane, Tim (2012). “What is the problem of non-existence?” In: *Philosophia* 40.3, pp. 417–434.

— (2013). *The objects of thought*. Oup Oxford.

Dancy, Russell Mather (2012). *Sense and contradiction: a study in Aristotle*. Springer Science & Business Media.

Davis, Wayne (2019). “Implicature”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2019. Metaphysics Research Lab, Stanford University.

Eklund, Matti (2019). “Fictionalism”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2019. Metaphysics Research Lab, Stanford University.

Field, Hartry (2008). *Saving truth from paradox*. Oxford University Press.

Findlay, JN (1963). “Meinong’s Theory of Objects and Values”. In:

Frege, Gottlob (1952a). “Function and Concept”. In: *Translations from the philosophical writings of Gottlob Frege*. Ed. by Peter Thomas Geach and Max Black. Basil Blackwell Oxford.

— (1952b). *Translations from the philosophical writings of Gottlob Frege*. Basil Blackwell Oxford.

Geach, Peter T (1965). “Assertion”. In: *The Philosophical Review* 74.4, pp. 449–465.

Gonzalez, Luis Estrada and Manuel Tapia-Navarro (2021). “" A Smack of Irrelevance" in Inconsistent Mathematics?” In: *The Australasian Journal of Logic* 18.5, pp. 503–523.

Green, Mitchell (2020). “Speech Acts”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2020. Metaphysics Research Lab, Stanford University.

Grice, H. P. (1968). “Utterer’s Meaning, Sentence Meaning, and Word Meaning”. In: *Foundations of Language* 4, pp. 1–18.

Grice, H Paul (1957). “Meaning”. In: *The philosophical review* 66.3, pp. 377–388.

Hansson, Sven Ove (2017). “Logic of Belief Revision”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2017. Metaphysics Research Lab, Stanford University.

Hawthorne, John (2006). *Metaphysical essays*. Oxford University Press on Demand.

Hodges, Wilfrid (2018). “Tarski’s Truth Definitions”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2018. Metaphysics Research Lab, Stanford University.

Hovda, Paul (2009). “What is classical mereology?” In: *Journal of Philosophical Logic* 38.1, pp. 55–82.

Husserl, Edmund (1901). *Logische untersuchungen*.

Hyde, Dominic (1997). “From heaps and gaps to heaps of gluts”. In: *Mind* 106.424, pp. 641–660.

Hyde, Dominic and Diana Raffman (2018). “Sorites Paradox”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2018. Metaphysics Research Lab, Stanford University.

Jacquette, Dale (2011). *Meinongian logic: The semantics of existence and nonexistence*. Vol. 11. Walter de Gruyter.

Kabay, Paul Douglas (2008). “A defense of trivialism”. PhD thesis.

Kripke, Saul (1976). “Outline of a theory of truth”. In: *The journal of philosophy* 72.19, pp. 690–716.

Künne, Wolfgang (2011). “Perception, Fiction, and Elliptical Speech”. In: *Thinking and the Structure of the World/Das Denken und die Struktur der Welt*. De Gruyter, pp. 259–267.

Lando, Giorgio (2017). *Mereology: A Philosophical Introduction*. Bloomsbury Press.

Lejewski, Czesław (1957). “Nicholas Rescher. Axioms for the part relation. Philosophical studies (Minneapolis), vol. 6 (1955), pp. 8–11.” In: *The Journal of Symbolic Logic* 22.2, pp. 213–214.

Lesniewski, Staislaw (1916). “Foundations of the General Theory of Sets (D.J. Barnett, Trans.)” In: *Collected works* 1, pp. 129–173.

Levi, Isaac (1978). “Subjunctives, dispositions and chances”. In: *Dispositions*. Springer, pp. 303–335.

Lewis, David (1982). “Logic for Equivocators”. In: *Nous*.

— (1991). *Parts of Classes*. Blackwell.

— (1993). “Mathematics in megethology”. In: *Philosophia Mathematica* 1.1, pp. 3–23.

Lewis, David (2004). “Letters to Priest and Beall”. In: *The Law of Non-Contradiction*. Ed. by J. C. Beall Graham Priest and Bradley Armour-Garb. Oxford: Oxford University Press, pp. 176–177.

Linnebo, Øystein (2017). “Plural Quantification”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2017. Metaphysics Research Lab, Stanford University.

Littmann, Greg and Keith Simmons (2004). “A Critique of Dialetheism”. In: *The Law of Non-Contradiction: New Philosophical Essays*. Ed. by Jc Beal G. Priest and B. Armour-Garb. Oxford, pp. 314–335.

Loss, Roberto (2018). “A Sudden Collapse to Nihilism”. In: *The Philosophical Quarterly* 68.271, pp. 370–375.

Lukasiewicz, Jan (1971). “On the principle of contradiction in Aristotle”. In: *The Review of Metaphysics*, pp. 485–509. Translated by V. Wedin.

MacBride, Fraser et al. (2020). “Frank Ramsey”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2020. Metaphysics Research Lab, Stanford University.

Mares, Edwin (2020). “Relevance Logic”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2020. Metaphysics Research Lab, Stanford University.

Meinong, Alexius (1960). “On the Theory of Objects”. In: *Realism and the Background of Phenomenology*. Ed. by Roderick Chisholm. Free Press, pp. 76–117.

Murzi, Julien and Massimiliano Carrara (2015). “Paradox and logical revision. A short introduction”. In: *Topoi* 34.1, pp. 7–14.

Myhill, John (1984). “Paradoxes”. In: *Foundations: Logic, Language, and Mathematics*. Springer, pp. 129–144.

Parsons, Terence (1990). “True contradictions”. In: *Canadian Journal of Philosophy* 20.3, pp. 335–353.

Pleitz, Martin (2018). *Logic, language, and the Liar paradox*. Mentis.

Popper, Karl R (1940). “What is dialectic?” In: *Mind* 49.196, pp. 403–426.

Priest, Graham (1979). “The logic of paradox”. In: *Journal of Philosophical logic* 8.1, pp. 219–241.

— (1994). “The structure of the paradoxes of self-reference”. In: *Mind* 103.409, pp. 25–34.

— (1999). “Perceiving contradictions”. In: *Australasian Journal of Philosophy* 77.4, pp. 439–446.

— (2002). *Beyond the limits of thought*. Oxford University Press.

— (2006a). *Doubt Truth to be a Liar*. Oxford University Press.

— (2006b). *In Contradiction*. Oxford University Press.

— (2010). “Inclosures, vagueness, and self-reference”. In: *Notre Dame Journal of Formal Logic* 51.1, pp. 69–84.

— (2014a). “Much Ado about Nothing”. In: *The Australasian Journal of Logic* 11.2.

— (2014b). *One: Being an Investigation into the Unity of Reality and of its Parts, including the Singular Object which is Nothingness*. Oxford University Press.

— (2016). *Towards non-Being: The Logic and Metaphysics of Intentionality*. Oxford University Press.

— (2017). "Entangled Gluons: Replies to Casati, Han, Kim, and Yagisawa". In: *International Journal of Philosophical Studies* 25.4, pp. 560–568.

— (2019). "It was so revolting I couldn't take my eyes off it". In: *Dialetheism and its Applications*. Springer, pp. 47–56.

Quine, Willard (2009). *Mathematical logic*. Harvard University Press.

Quine, Willard V (1948). "On What There Is". In: *The Review of Metaphysics* 2.1, pp. 21–38.

Quine, Willard Van Orman (2011). *On what there is*. Princeton University Press.

Ramsey, Frank Plumpton (1926). "The foundations of mathematics". In: *Proceedings of the London Mathematical Society* 2.1, pp. 338–384.

Reicher, Maria (2019). "Nonexistent Objects". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2019. Metaphysics Research Lab, Stanford University.

Rescher, Nicholas and Robert Brandom (1980). "The logic of inconsistency". In:

Ripley, David (2011). "Negation, denial, and rejection". In: *Philosophy Compass* 6.9, pp. 622–629.

— (2013). "Paradoxes and failures of cut". In: *Australasian Journal of Philosophy* 91.1, pp. 139–164.

Rosen, Gideon and Cian Dorr (2002). "Composition as a Fiction". In: *The Blackwell guide to metaphysics*, pp. 151–174.

Rossberg, Marcus (2013). "Too good to be "just true"". In: *Thought: A Journal of Philosophy* 2.1, pp. 1–8.

Routley, Richard (1979). "Dialectical logic, semantics and metamathematics". In: *Erkenntnis*, pp. 301–331.

— (1980). *Exploring Meinong's jungle and beyond: an investigation of noneism and the theory of items*. Citeseer.

— (2019). *Ultralogic as universal?* Springer.

Routley, Richard and Robert K Meyer (1976). "Dialectical logic, classical logic, and the consistency of the world". In: *Studies in Soviet Thought*, pp. 1–25.

Sainsbury, Richard Mark (2009). *Paradoxes*. Cambridge University Press.

Sgaravatti, Daniele and Giuseppe Spolaore (2018). "Out of Nothing". In: *Thought: A Journal of Philosophy* 7.2, pp. 132–138.

Shapiro, Stewart (2004). "Simple truth, contradiction, and consistency". In: *The Law of Non-Contradiction*. Ed. by J. C. Beall Graham Priest and Bradley Armour-Garb. Oxford: Oxford University Press, pp. 336–354.

Sider, Theodore (2014). "Consequences of Collapse". In: *Composition as Identity*. Ed. by Aaron J. Cotnoir & Donald L.M. Baxter. Oxford University Press.

Sider, Theodore et al. (2001). *Four-dimensionalism: An ontology of persistence and time*. Oxford University Press on Demand.

Simons, Peter (1987). *Parts: A study in ontology*.

Smith, Nicholas JJ (2000). "The principle of uniform solution (of the paradoxes of self-reference)". In: *Mind*, pp. 117–122.

Sorensen, Roy (2003). *A Brief History of the Paradox: Philosophy and the Labyrinths of the Mind*. Oxford University Press.

— (2018). "Vagueness". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2018. Metaphysics Research Lab, Stanford University.

— (2020). "Epistemic Paradoxes". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2020. Metaphysics Research Lab, Stanford University.

Tanaka, Koji (2005). "The AGM theory and inconsistent belief change". In: *Logique et analyse*, pp. 113–150.

Tarski, Alfred (1935). "Zur Grundlegung der Boole'schen Algebra. I". In: *Fundamenta mathematicae* 1.24, pp. 177–198.

— (1936). "Der Wahrheitsbegriff in den formalisierten Sprachen". In:

— (1956). "The concept of truth in formalized languages". In: *Logic, semantics, metamathematics* 2, pp. 152–278.

Uzquiano, Gabriel (2014). "Mereology and modality". In: *Mereology and location*. Ed. by Shieva Kleinschmidt. Oxford University Press, pp. 33–56.

Van Inwagen, Peter and Peter Van Inwagen (1990). *Material beings*. Cornell University Press.

Varzi, Achille (2019). "Mereology". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2019. Metaphysics Research Lab, Stanford University.

Weber, Zach (2010). "A paraconsistent model of vagueness". In: *Mind* 119.476, pp. 1025–1045.

— (2012). "Transfinite Cardinals in Paraconsistent Set Theory." In: *Review of Symbolic Logic* 5.2.

— (2021). *Paradoxes and Inconsistent Mathematics*. Cambridge University Press.

Weber, Zach and Aaron J Cotnoir (2015). "Inconsistent boundaries". In: *Synthese* 192.5, pp. 1267–1294.

Wittgenstein, Ludwig (1956). *Remarks on the Foundations of Mathematics*. Balckwell.

Yagisawa, Takashi (2017). "Gluons of Different Colors". In: *International Journal of Philosophical Studies* 25.4, pp. 555–559.

Young, Gareth (2015). "Shrieking, just false and exclusion". In: *Thought: A Journal of Philosophy* 4.4, pp. 269–276.

Zalta, Edward (2012). *Abstract objects: An introduction to axiomatic metaphysics*. Vol. 160. Springer Science & Business Media.

Zolghadr, Behnam (2019). "Being and Nothingness". In: *The Australasian Journal of Logic* 16.3, pp. 68–82.