University of Alberta

The false only problem for dialetheism

by

Nika Pona

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Arts

Department of Philosophy

Fall 2013
Edmonton, Alberta

To my mother

Abstract

Paraconsistent logics reject the validity of the classical principle of *Explosion* that says that from a contradiction one can derive any proposition. A *dialetheist* is a paraconsistent logician who claims that this principle is invalid, because there are statements that are *both true and false*. The paradigmatic example of such contradiction is the Liar Paradox – informally, a sentence that says of itself: “I am not true”. By the dialetheists’ view, philosophers and logicians for centuries tried to find a consistent solution to this paradox in vain. The paraconsistent solution is to change the classical logic to a paraconsistent logic and accept that the Liar is *both true and false*.

In this thesis I will discuss an objection to the claim that one would be better off if she switched to the dialetheic paraconsistent logic. The problem is that the dialetheist can’t express the familiar notions of truth and falsity *simpliciter*. That is, she can’t describe the consistent domains, neither can she reason about them.

Acknowledgements

I am grateful to the amazing community of faculty, graduate students and support staff of the Department of Philosophy for providing me with such a wonderful intellectual experience during the past two years.

The courses of Bernie Linsky, Jeff Pelletier, Kata Bimbo and Sam Hillier, as well the the weekly discussions of the Logic Reading Group shaped my interests in logic and philosophy and gave me tools and knowledge necessary to develop these interests. I am particularly grateful to Allen Hazen and Jeff Pelletier for helping me with the thesis project.

Special thanks to my fellow graduate students for passionate philosophical discussions outside the classroom and to Sussanne McDonald who made my encounters with academic bureaucracy as painless as possible.

Finally, thanks to Pierre Deffontaines for taking care of me while I was writing this thesis.

Table of Contents

Introduction

Paraconsistent logics reject the validity of the principle of *Explosion*, $\neg A, A \models B$, that says that from a contradiction one can derive any proposition. A *dialetheist* is a paraconsistent logician who claims that there is at least one model where $\neg A$ and A are true and B is false, i.e., that there are true contradictions. The paradigmatic example of such contradiction is the Liar Paradox that arises when one adds a truth-predicate $T(x)$ to the language, informally, this is a sentence that says of itself: "I am not true" ($A \leftrightarrow \neg T(A)$). By the dialetheists' view, philosophers and logicians for centuries tried to find a consistent solution to this paradox in vain. Their solution is to change the classical logic that is intolerant to inconsistencies (this is the fact expressed by the validity of *Explosion*) to a paraconsistent logic and accept that the Liar is both true and false. This logic is supposed to be more universal, since it could integrate the paradoxical cases with the familiar, classical, reasoning.

In my thesis I will discuss an objection to the claim that one doesn't lose anything important from the classical logic when switching to a paraconsistent logic. Namely, the objection goes, the dialetheist can't express the familiar notions of truth and falsity *simpliciter*. That is, she can't describe the consistent domains, neither can she reason about them. (E.g., given the paraconsistent semantics for negation she is not able to say that some proposition A is *not* *both true and false.*)

Here I will analyze this objection as mounted against a particular paraconsistent logic LP and its improved version $LP^{\rightarrow}$. $LP^{\rightarrow}$ is LP – a three-valued logic dual to the Strong Kleene paracomplete logic K_3 (a logic that allows for the truth-value gaps, a model where neither A nor $\neg A$ is true) – equipped with an "appropriate" conditional, a conditional that validates most of the important principles, like *Modus Ponens* $A,\ A \supset B \models B$ and *Modus Tollens* $\neg B, A \supset B \models \neg A$, and avoids the Curry Paradox (informally, a sentence that says of itself: "If I am true, then anything is true", $T(A) \leftrightarrow (T(A) \rightarrow T(B))$)[1].

The discussion will proceed as follows:

In the *Preliminaries* I will give the semantics for LP, leaving the definitions of "appropriate conditionals" and $LP^{\rightarrow}$ for Chapter 2, where they will be treated in detail.

Chapter 1 will introduce the definition of the *false only* problem, as well as an explanation of its importance. In addition, I will go through some familiar "object language" solutions that consist in introducing a *false only* operator or predicate to the object language of LP or $LP^{\rightarrow}$. They are unsuccessful, since they give rise to new paradoxes ("extended Liars" or "strengthened Liars").

In *Chapter 2* I will argue, contrary to the opinion established in dialetheic circles, that there is one successful object language solution to the *false only* problem. Namely, one could express that some statement A is *false only* by *falsum*-negating it, i.e., asserting $A \rightarrow \perp$ (informally, "If A is true, then anything follows"). This solution avoids extended Liars, since the principle

[1]If A is true, i.e., $T(A)$, then one gets to $T(B)$ by applying *Contraction* and *Modus Ponens*. If A is false, then $T(A) \rightarrow T(B)$ has to be false as well (by equivalence), that is $T(B)$ has to be false and $T(A)$ true. Thus, A can't be false on pain of contradiction, but can't be true either, on pain of triviality.

of *Contraction* ($A \to (A \to B) \models A \to B$) is invalid in $LP^{\to}$, as well as the Law of Excluded Middle for the *falsum*-negation ($(A \to \perp) \vee A$). On the other hand, the conditional "If A is true, then triviality follows" doesn't appear to express adequately that A is *false only*, that is, that A is to be denied[2]. I will claim that the appearance is misleading in this case and argue that there is a way to interpret the conditional in the *falsum*-negation that makes it a plausible account of the falsity *simpliciter*.

In the final *Chapter 3* I will discuss an alternative to the "object language" solutions and show that it faces serious problems. In particular, appealing to the speech-act theory and treating *false only* as a force operator of *denial* on the sentences' contents (like command, question, etc.) suffers from what is familiar in the philosophy of language and moral philosophy as the Frege-Geach problem. It amounts to showing the impossibility of dividing the language into the two parts, where the first one contains the terms that have their meaning specified by truth-conditions and other – by force operators. If one were to accept such a view, the objection goes, then one is to deny the possibility of using the second category's terms in the complex sentences. It is highly counter-intuitive that one can't use "assert" and "deny" in a conditional, for instance ("If one denies A, then one denies B").

I will conclude by arguing in favor of the *falsum*-negation and providing the directions for further investigation.

[2]In the dialetheic framework we are not supposed to deny claims that are *at least false* as the Liar sentence, but only the one that are false *simpliciter*

Preliminaries: Semantics for LP

In this thesis when I say "paraconsistent logic" I mean the logic LP and "paraconsistent logic with an appropriate conditional" is what I call $LP^{\rightarrow}$. Here is the definition for LP:

LP[3]

The syntax is that of classical propositional logic with connectives $\{\wedge, \vee, \neg, \supset\}$; equivalence $A \leftrightarrow B$ can be defined as usual: $(A \supset B) \wedge (B \supset A)$.
For semantics, let $\langle \mathcal{V}, \mathcal{D}, \{f_c; c \in \mathcal{C}\} \rangle$ be the structure that defines the logic LP. $\mathcal{V}$ is then the set of truth values with n members, $1 \leq n$. $\mathcal{D}$ is a set of *designated values*, the values that are preserved in valid inferences, $\mathcal{D} \subseteq \mathcal{V}$. In classical logic, for instance, the only designated value is 1 or *true*. $\mathcal{C}$ is the set of connectives; for every connective c, f_c is the truth function corresponding to it. For example, in classical logic $\neg$ denotes 1-place function $f_\neg$, such that $f_\neg(0) = 1$ and $f_\neg(1) = 0$. An interpretation v is a map from propositional parameters to V. If c is an n-place connective, $v(c(A_1, ..., A_n)) = f_c(v(A_1), ..., v(A_n))$. An inference from a set of premisses Σ to the conclusion A is semantically valid, $\Sigma \models A$, if and only if there is no interpretation v, such that for all $B \in \Sigma$, $v(B) \in \mathcal{D}$ and $v(A) \notin \mathcal{D}$. A is a logical truth if and only if $\emptyset \models A$, i.e., for

[3] [Priest 2006, 18.2; Priest 2007, 7.2-7.5]

every interpretation $v(A) \in \mathcal{D}$[4].

LP is a logic with $\mathcal{V} = \{1, i, 0\}$, where 1 and 0 are thought of as *true* and *false* respectively, i – as *both true and false*. $\mathcal{D} = \{1, i\}$[5]. Functions f_c corresponding to the connectives of *LP* are the following:

$f_\neg$	
1	0
i	i
0	1

$f_\wedge$	1	i	0
1	1	i	0
i	i	i	0
0	0	0	0

$f_\vee$	1	i	0
1	1	1	1
i	1	i	i
0	1	i	0

$f_\supset$	1	i	0
1	1	i	0
i	1	i	i
0	1	1	1

This logic is paraconsistent, because $A \wedge \neg A \not\models B$, i.e., the principle of *Explosion* or *Ex Falso Quodlibet* is not valid. Take $v(A) = i$, then $v(\neg A) = i$, and $v(B) = 0$: all premises have designated values, while the consequence has not. *Modus Ponens* doesn't hold either, $A, A \supset B \not\models B$. Again, take $v(A) = i$ and $v(B) = 0$. $v(A \supset B) = i$, by the $\supset$ truth-table, so all the premises have the designated value, but not the consequence.

Another way of thinking about the evaluation v is to take it to be a subset of $F \times \{1, 0\}$ ("relational semantics"), where F is a set of formulas of *LP* that are related to either 1, 0 or *both*, which equivalent to interpreting it as an evaluation function that maps a formula to the set $\mathcal{V}$ of truth-values to which the formula is related, $\mathcal{V} = \{\{0\}, \{1\}, \{0, 1\}\}$. The clauses for the connectives in this case would be the following [Priest 2006, Ch. 18.3; 2002, pp. 308-310]:

$1 \in v(\neg A)$ iff $0 \in v(A)$ and $0 \in v(\neg A)$ iff $1 \in v(A)$[6]

[4]I follow Priest and Beall in assuming that the classical reasoning is acceptable in the metatheoretic context [Priest 2007, pp. 584-585; Priest 2006, Ch. 18.5; Beall 2009, Ch. 1].

[5]The same logic with $\mathcal{D} = \{1\}$ is Strong Klenee three-valued logic K_3.

[6]Note the difference between this clause (De Morgan negation) and the classical one (Boolean negation): $1 \in v(\neg A)$ iff 1 not $\in v(A)$ and $0 \in v(\neg A)$ iff $1 \in v(A)$. Here from the assumption that $\{1, 0\} \in v(\neg A)$ we arrive at the consequence that $1 \in$ and $\notin v(A)$, which is not acceptable given that metatheory is supposed to be consistent[Priest 2002, p. 385; 2008, Ch. 5].

$1 \in v(A \wedge B)$ iff $1 \in v(A)$ and $1 \in v(B)$

$0 \in v(A \wedge B)$ iff $0 \in v(A)$ or $0 \in v(B)$

$1 \in v(A \vee B)$ iff $1 \in (A)$ or $1 \in v(B)$

$0 \in v(A \vee B)$ iff $0 \in (A)$ and $0 \in v(B)$[7]

$\Sigma \models A$ iff for any v if $1 \in v(B)$ for all $B \in \Sigma$, then $1 \in v(A)$

$\models A$ iff for any v $1 \in v(A)$.

The invalidity of *Explosion* would be shown then as follows. $A \wedge \neg A \not\models B$: Take $\{1,0\} \in v(A)$ and $\{0\} \in v(B)$. Then $\{1,0\} \in v(\neg A)$ by the negation clause, so is $A \wedge \neg A$ by the conjunction clauses. Then $A \wedge \neg A \not\models B$, because 1 is in the evaluation of every premise, but not in the conclusion's; and similarly for *Modus Ponens*.

This modification of the formalism will be important in the discussion of the Extended Liars (Chapter 1).

First-order LP is characterised by a language extended with predicate symbols and a structure of the form $\langle D, d, \mathcal{V}, \mathcal{D}, \{F_c : c \in \mathcal{C}\}\rangle$, D is a non-empty domain, d maps every constant into D, and if P is an n-place predicate, d maps P to a function from n-tuples of the domain into the set of truth-values. For atomic sentences, $Pc_1...c_n$, $v(Pc_l...c_n) = d(P)(d(c_l)...d(c_n))$ [8]. $v(\forall x A)$ is defined as the greatest lower bound of $\{v(A_x(k_d)) : d \in D\}$, $\exists x$ – as the least upper bound, when ordering on $\mathcal{V}$ is $1 < i < 0$, and the latter means the substitution of named objects from D in Ax.

In terms of relational semantics:

An interpretation $\mathcal{I}$ is a pair $\langle D, d\rangle$, where D is the domain of quantifica-

[7] $A \supset B$ is defined as $\neg A \vee B$.

[8] For an alternative presentation in terms of relations see [Priest 2006, Ch. 5.3.; for a quantified logics, see 2002, pp. 332-335]

tion, d maps every constant into D, and if P is an n-place predicate, d maps P into a pair $\langle P^+, P^- \rangle$, $P^+ \cup P^- = D^n$. P^+ is an *extension* and P^- *anti-extension* of a predicate P. Truth values for atomic sentences are assigned in the following way:

$1 \in v(Pc_1...c_n)$ iff $\langle d(c_1), ..., d(c_n) \rangle \in P^+$

$0 \in v(Pc_1...c_n)$ iff $\langle d(c_1), ..., d(c_n) \rangle \in P^-$

The quantifiers:

$1 \in v(\forall xA)$ iff for every $c \in D, 1 \in v(A(x/c))$

$0 \in v(\forall xA)$ iff for some $c \in D, 0 \in v(A(x/c))$

$1 \in v(\exists xA)$ iff for some $c \in D, 1 \in v(A(x/c))$

$0 \in v(\exists xA)$ iff for every $c \in D, 0 \in v(A(x/c))$.

We can add an identity predicate to the language as well $v(=)(d_1, d_2) \in \mathcal{D}$ iff $d_1 = d_2$, as well as a Truth-predicate: $1 \in v(A)$ iff $A \in T^+$, and a False-predicate $1 \in v(\neg A)$ iff $A \in F^+$, where A is a closed sentence. The Truth predicate satisfies the Truth-scheme $T(A) \leftrightarrow A$, given the "appropriate conditional" of the logic $LP^{\rightarrow}$[9]. We require also that If $0 \in v(A)$, then $A \in T^-$, although not the converse. For the sake of convenience, I will specify the semantics of $LP^{\rightarrow}$ in the *Chapter 2*.

[9]One could think of the Truth-scheme in the rule form as well: $T(A) \models A$ and $A \models T(A)$, in order to avoid appeal to conditionals. Then it is easy to check that those are the valid inferences in the logic specified. $1 \in v(T(A))$ means that $A \in T^+$, by meaning of satisfaction of predicates. Then, $1 \in v(A)$, by the definition of Truth predicate. If $0 \in v(T(A))$, then $1 \in v(\neg(T(A)))$, i.e. $1 \in v(F(A))$, which means that $A \in F^+$ and then $1 \in v(\neg A)$. By the negation clause, we have: $0 \in v(A)$. Thus, there no evaluation where $1 \in v(T(A))$, but $1 \notin v(A)$. Similarly for other direction.

Chapter 1

The *false only* problem for LP

1..1 What is the *false only* problem?

The *false only* problem for dialetheist – one who finds there to be true contradictions – amounts to inability of her favourite paraconsistent logic (here: LP and $LP^{\rightarrow}$) to express that some statements are *false only*, that is, that they are *not dialetheias*[1]. It is a *problem*, because it appears that the norms of rational discourse require that we could say that some claims are *definitely false* and some – *simply true*, i.e., that some claims are *not dialetheias*[2].

In LP, the most natural way to express that A is *false only* seems to be to assert $F(A) \wedge \neg T(A)$ ("A is *false* and *not true*"). By definition of *not true*[3],

[1]It is also known as the *true only* problem, for dialetheist can't express that something is *true simpliciter* either. The solution to the *true only* problem will stem from the solution to the *false only* problem, but since the latter seems to be more interesting due to its relation to disagreement, I will concentrate on it. Throughout my discussion I will use terms *false only*, *just false*, *definitely false*, *simply false* and *false simpliciter* to name the same issue.

[2]The *false only* problem frequently serves as a criticism (e.g., [Shapiro 2004]), but it is also the worry dialetheists have themselves: "That some sentences are true and false is one thing; however, the dialetheic position is rational only if at least some sentences are *just true*. The worry is whether the dialetheist can give an adequate account of "just true" without the position exploding into triviality" [Beall 2007, p. 5].

[3]Typically, *falsity* and *not truth* are taken to be equivalent. [Priest 2006] challenges this equivalence: he rejects $F(A) \rightarrow \neg T(A)$, since he wants to say that the Liar is *true*

this conjunction becomes $F(A) \wedge F(A)$, that is, $F(A)$, by simplification. And $F(A)$ can't express *falsity simpliciter*, since then the Liar sentence that is *both true and false* is *false only* as well. Given that we want the notion *false only* to describe only *non-dialetheias*, i.e., to exclude that the *simply false* statement is *true* as well, we can't accept such a conclusion.

Maybe we could indicate that A is *false only* by saying that it is *false* ($F(A)$) and *not a dialetheia*? Since it is straightforward how to say that a sentence, A, is a *dialetheia*[4] (assert $T(A) \wedge F(A)$), one might suspect that the negation of this statement would solve our problem. Unfortunately, it doesn't, due to LP's semantics of negation [see p. 5 of the present work]. If $T(A) \wedge F(A)$ expresses that A is a *dialetheia*, its negation will express the same thing, since the negation of a dialetheia is *both true and false* too. Moreover, $\neg(A \wedge \neg A)$ is a theorem of LP, which means that, for every A, $\neg(A \wedge \neg A)$ is *true* or *both true and false*.

It may appear that this criticism begs the question, since dialetheist doesn't have to give up on these definitions of *truth* and *falsity simpliciter* if it is shown that they are inconsistent, since she already accepts some contradictions. In order to avoid such an accusation, one has to show why dialetheist can't accept this contradiction on pain of some bad consequences for dialetheist, where

and *false*, but not *true* and *not true*. He seems to be an exception to the general trend to preserve the transparency of truth predicate that results in equivalence of *falsity* and *not truth* [cf. Beall 2009; Ripley 2011a; 2013; Field 2008]. Nothing in the following should hang on this difference, since the right-left direction of the equivalence ($\neg T(A) \rightarrow F(A)$) stays unchallenged. (You can think of this equivalence in a rule form, if you are worried about the conditional.)

[4]The dialetheists agree that the Liar paradox is a clear example of a sentence that is *both true and false*. Priest [2013] argues for dialetheism as a uniform solution to Sorites, Russell's and Curry paradoxes. Sometimes the cases of inconsistent laws or moral obligations are considered to be dialetheias as well [Priest 2002, p. 127-129; Bremer 2005]. Some argue that there are real world examples of dialetheias: Priest [2006, Ch. 11] takes change and motion to be the prime example of real world dialetheias (hereby suggesting the dialetheic solution to Zeno's paradoxes).

"bad" is something the latter doesn't accept or *shouldn't* accept. The *false only* problem is "bad" for dialetheist in both ways.

First, the dialetheic solution to the Liar paradox (accepting that the Liar sentence is *both true and false*) advertises "non-artificiality" and the lack of "expressibility problem" as its advantage – here the *false only* problem is used to show that this is false advertising. In the dialetheic view, the solution proposed by type-theorist[5] doesn't explain how the language really works ("artificiality"); furthermore, it suffers from the same problem as the gap-theorist's solution[6] – both solutions make use of the concepts that don't make sense according to these solutions. For instance, the gap theorist won't be able to formalize her meta-theoretic notion *not true* on pain of contradiction (the "expressibility problem"). The type-theorist in turn wouldn't be able to state what her theory amounts to[7], since this statement would have to appeal to quantification on hierarchies prohibited by type-theory[8]. In the same vein, one could engage into *tu quoque* criticism and say that dialetheist can do no better, since she can't express the notions she uses in her (consistent) meta-theory. Usually, that's what "strengthened Liars" aim at[9].

[5]Roughly, in order to avoid self-reference of the Truth-predicate that leads to paradox, one postulates that there are is a hierarchy of Truth-predicates: the Liar sentence won't be well-formed, since one can't apply the Truth-predicate to the sentence of the same "type" (level).

[6]Gap-theorist is someone who suggests to solve the Liar paradox by saying that the Liar sentence is *neither true nor false*.

[7]Informally, "The Truth predicate of level n is defined at level $n+1$".

[8][Fitch 1946] calls pointing out to such fact an *ad hominem* criticism, since what is involved in such views (like type theory) is not a contradiction, but a kind of "self-referential incoherency", when one does something she is not supposed to do. He considers this to be a serious philosophical criticism. [Bremer 2005, pp. 27-30, pp.141-142] points in the similar direction.

[9]The strengthened, or extended Liar is a paradox that arises from the solution to the original Liar paradox. For instance, if you think that one could solve the Liar paradox by saying that the Liar sentence is *neither true nor false*, consider a sentence A: "This sentence is *neither true nor false* or *false*". If this sentence A is *true*, then it is *neither true nor false* or *false*, which means that it is *not true*. If A is *not true*, then what is says is correct (it is

Although such a criticism may have success in persuading the opponent to revise her position or her criticism, it doesn't help in deciding which logic to choose, which one is correct, better, more practical – whatever one takes to be the value of a logic. It seems to be stronger (and more interesting) to argue that not only is the dialetheist in a kind of self-incoherent position (there is always a possibility that she could accept such a state[10]), but that her views lead to some generally unpleasant consequences that neither she, not others can accept. In such a case the criticism serves not only to undermine dialetheic views, but also to advance the discussion about what one should have or avoid in a logical theory.

In this thesis I will talk about one such bad consequence for the dialetheist, namely, her inability to express genuine disagreement.

Before explaining what this amounts to I will mention another unwanted consequence that arises from the dialetheist's constraints on expressibility. The dialetheic treatment, or lack thereof, of *simple falsity* poses a problem if one wants to form an interesting conditional using dialetheias. For instance, consider a conditional: "If the Liar sentence is *both true and false*, then dialetheism is *true*" (1)[11]. We can argue on grounds independent from the Liar's dialetheic status that the dialetheism, δ, is incorrect, i.e., *simply false* (2) (that's how the debate around the *Law of Non-Contradiction* usually proceeds) and so con-

---

either *neither true nor false* or *false*), which means that it is *true*. In both cases we arrive at a contradiction: A is *true* and *not true*. One could try avoiding this contradiction by saying that it doesn't follow from the fact that a sentence is *neither true nor false* or *false* that it is *not true* [see Beall 2006; 2007].

[10] In the *Conclusion* I will mention the "paracoherentism" – a view contrary to Fitch's, that self-incoherency is not problematic.

[11] [Shapiro 2004] talks about the similar issue for dialetheist speech act approach to denial: the speech act explanation is not very useful, since one can't appeal to speech acts when talking about hypotheses, but we would want to be able to formalize the hypotheses in our logic.

clude that one should not consider Liar, λ, as *both true and false* (4). However, by the semantics for negation in LP [see page 5 of the present work], $\neg\lambda$ will be *both true and false* as well[12].

1. $(T(\lambda) \wedge \neg T(\lambda)) \rightarrow T(\delta)$

2. $\neg T(\delta)$

3. $\neg T(\delta) \rightarrow \neg(T(\lambda) \wedge \neg T(\lambda))$ (1, *Contraposition*)

4. $\neg(T(\lambda) \wedge \neg T(\lambda))$ (2, 3, *Modus Ponens*[13].)

So, from disproving the dialetheism as a theory we don't get that the Liar is *not a dialetheia*, but its negation that is *both true and false*. This criticism might not be very convincing for a dialetheist, since she can accept further contradictions. Moreover, the validity of *Contraposition* may be challenged by some dialetheists [Priest 2006, Ch. 6]. The problem with such a conclusion is not that it is incoherent in some sense, but that it seems that it prevents us from making an interesting conditional involving dialetheias – a conditional that would require formulating the idea that something can be shown to be not a dialetheia.

[12] Here I use the dialetheism for the sake of example. We could talk about any kind of conditionals involving contradictions.

[13] What if *Modus Ponens* is invalid as [Beall 2013a] thinks? In his view, this inference is valid: (*) $\neg T(\delta) \rightarrow \neg(T(\lambda) \wedge \neg T(\lambda)), \neg T(\delta) \models \neg(T(\lambda) \wedge \neg T(\lambda)), T(\delta) \wedge \neg T(\delta)$. By Beall's view, what the invalidity of *Modus Ponens* means is that we are bound to use "extra-logical" means to decide what inference to accept. This explains why *Modus Ponens* *seems* to be valid: usually we don't choose a contradiction on the right hand side of the turnstile among the two disjuncts that follow from the premisses, because in most cases it is not a rational thing to do. It is rational only when we have a true contradiction in the premises. This explanation doesn't help in our (*) case though. By this logic we have to rationally choose between these alternatives: either we accept $\neg(T(\lambda) \wedge \neg T(\lambda))$ as a true contradiction, which leaves us with a true contradiction and false theory that says that there are true contradictions, or we accept that dialetheism is both true and false.

1..2 The genuine disagreement

The *false only* problem is damaging, because it is equivalent to one of the most commonly accepted objections to the dialetheism: that if somebody were to accept a theory that there are true contradictions, it would be impossible to argue with this person, because she wouldn't be able to express or even make sense of the *genuine disagreement*. In a paraconsistent setting one can't express that some statements are *false only* and therefore participate in the rational dialogue. This criticism is stronger than *ad hominem*, because it shows not only that dialetheist can't do what she is supposed to do better, but that what she can't do is a serious problem in general.

The term "genuine disagreement" points to the relation of the linguistic acts of assertion and denial (agreement and disagreement) to rationality (acceptance and rejection). Roughly, when a person expresses a genuine disagreement, she intends her audience to understand that her position is incompatible with the listener's and that the latter has to revise his beliefs[14].

[14]One might say that this definition doesn't capture the meaning of the term to its full extent, for it is common for people to disagree with somebody with the intention to make a correction to their interlocutors' views, and not to point out to the incompatibility of their positions or the necessity of the radical belief revision. (Consider the sentence: "It is not *a* beer – it is the best beer in the world!". The second part of the sentence specifies the first one, rather than contradicts it). One could argue that the genuine disagreement is the only kind of disagreement available despite the appearances and that in the given counter-example the incompatibility of the two statements is presupposed, that this incompatibility is implicit there and can be always made explicit by adding the "missing statements", analogous to an enthymematic argument ("It is not *a* beer (if by *a* beer you mean an *ordinary* beer) – it is the best beer in the world!"). Alternatively, one could take these counter-examples at face value and draw a distinction between the "genuine disagreement" (involving incompatibility relation between the statements) and the "corrective disagreement" (without such a relation, as in the beer example). Either way works well for this chapter's discussion, but I prefer to use the latter as I am interested in the clear-cut cases of the explicit, i.e., genuine, disagreement. For this reason in this chapter I am using the terms "disagreement" and "genuine disagreement" interchangeably. The examples of the "corrective disagreement" will show up in the discussion of denial and assertion, where they contribute to establishing that the assertion of the negation of a statement is not equivalent to the denial of that statement, so this notion will be useful as well

With the dialetheist you may find it difficult to interpret the situation when in response to your claim that A, you hear "$\neg A$" from your dialetheist friend. Given that your friend may accept the claims of the form "$A \wedge \neg A$", $\neg A$ doesn't have to mean that she disagrees with you (i.e., wants you to discard your belief that A), as she might just be pointing out to the relevant piece of information $\neg A$ that is true as well as A. In such situation you can't tell if the dialetheist is expressing genuine disagreement or corrective disagreement (see the footnote #13: the latter doesn't have a purpose of cancelling the statement it is addressed at; rather, it adds more information). Further, it might be the case that not only that the dialetheist failed to express her attitude in an unambiguous way, but that she wasn't even able to have this intention (to express the genuine disagreement), for this notion (involving the incompatibility relation[15]) doesn't make sense in her language. Moreover, when the dialetheist says "A" and you think that A *is false* and $\neg A$ *is true*, i.e., that she is wrong in her beliefs, asserting $\neg A$ won't achieve your intended

[15]Incompatibility is the basis of the Sheffer stroke $A \uparrow B$ that expresses that A and B both can't be true. Would it allow the dialetheist to express genuine disagreement with A, i.e., that A is *false only?*. The truth-conditions of the Sheffer stroke are the following: $1 \in v(A \uparrow B)$ iff $0 \in v(A)$ or $0 \in v(B)$; $0 \in v(A \uparrow B)$ iff $1 \in v(A)$ and $1 \in v(B)$. Let's see if we could formulate an Extended Liar using the Sheffer stroke to express that L is *false only*. Consider the sentence $L \leftrightarrow (F(L) \wedge (T(L) \uparrow F(L)))$ ("L is *false* and it can't be both *true* and *false*"). If L is *true*, $T(L)$, i.e., $1 \in v(L)$, then $1 \in v(F(L) \wedge (T(L) \uparrow F(L)))$. This is the case iff $1 \in v(F(L))$ and $1 \in v(T(L) \uparrow F(L))$. Since $1 \in v(F(L))$, it has to be the case that $0 \in v(T(L))$ for $(T(L) \uparrow F(L))$ to be true. Thus we derive $\neg T(L)$ from the assumption that $T(L)$. If we assume that L is *false*, $F(L)$, i.e., $0 \in v(L)$, then $0 \in v(F(L) \wedge (T(L) \uparrow F(L)))$. It means that either $0 \in v(F(L))$ or $0 \in v(T(L) \uparrow F(L))$. By assumption, $1 \in F(L)$, so we are left with one case when the disjunciton $0 \in v(F(L))$ or $0 \in v(T(L) \uparrow F(L))$ is *true* and doesn't immediately lead to a contradiction. If $1 \in v(F(L))$ and $0 \in v(T(L) \uparrow F(L))$, then it has to be that $1 \in v(T(L))$, contrary to the assumption that $F(L)$. In all the cases we arrive at the contradiction that $T(L)$ and $\neg T(L)$ (or at $T(L)$ and $F(L)$ if one takes *falsity* and *not truth* not to be equivalent [see the footnote # 3 on p. 8 of the present work]). In a paraconsistent setting we could accept this contradiction or the fact that L is *true* and *false*, as we did in the case of the original Liar paradox. We could say that L is *both true and false*, but in this case the aim of expressing that L is *false only* is not achieved. This result is the same as in the case of introducing explicit *false only* operators (see below). Thanks to B. Linsky for bringing up this question.

goal, because the dialetheist might think that $\neg A$ *is true* as well. Thus, you are in a position where neither you nor your opponent can express disagreement. What both of you would like to convey in this situation is that A *is not true* and $\neg A$ *is true*, but *not both* A *and* $\neg A$, i.e., that A is *false only.* But asserting it as it stands won't do the required job, as the dialetheist can accept a further contradiction (A *is not true* and $\neg A$ *is true*, but *not both* AND A *is true*, for instance).

If you accept the paraconsistent rules of the game, you end up in a situation where you can't convey disagreement to your interlocutor using the "ordinary means" ($\neg A$)[16], as in the paraconsistent context they don't presuppose the incompatibility needed for the genuine disagreement. The question is, if there are other means to achieve this goal. In my thesis I will analyze the positive answer to this question.

If the answer to this question (are there ways to genuinely disagree in the dialetheist framework?) is "no", many[17] would consider this as a serious criticism of the dialetheic position, as it amounts to more than the practical problems with respect to the conversation ("well, it is tricky to talk to a dialetheist, so what"). Genuine disagreement is more important than that as it is considered to be the basis of the rational dialogue, which in its turn governs, that is, puts in motion the belief revision process, necessary for an agent to count as rational. If the acts of expressing agreement or disagreement on the linguistic level correspond to the cognitive acts of rational acceptance and rejection (understood as proportional to the amount evidential support available and accordance to the basic cognitive values, such as simplicity, explanatory

[16] $\neg$ is abbreviating "not", "no", "it is not the case that", etc.

[17] See contributions to [Armour-Garb, Beall, Priest 2004] by Grim, Sainsbury, Lewis, Shapiro; [Priest 2006, Ch. 20; 2008, Ch. 7-8] discusses a similar worry.

power, productivity, etc.[18]), then not being able to disagree is quite a serious problem, as it means that one is not rationally rejecting anything[19], which is far from being a rational position.

One might think that adding an "explosive" or "incompatibility" operator ("true only", "false only") would solve the problem. In this chapter I will show that such a solution easily leads to the Extended Liars an so triviality, unless one gives up on certain intuitive principles.

1..3 Object language solutions to the *false only problem*

Here I will discuss "object language" solutions to the *false only* problem. These solutions involve defining a *just false* predicate or operator using the formalism available in LP. They are usually followed by the "extended Liars" that show that one can't define such a predicate on pain of triviality. Although usually extended Liars serve as *tu quoque* arguments that the dialetheic solution suffers from the same problems as other solutions, these attempts are important for my further discussion, for they show the problems related to "object language" solutions to the *false only* problem. I will discuss only some of them, for the sake of illustration, since it is a settled point in the literature that one can't define a *false only* predicate without falling into triviality [e.g., see Beall 2009, Ch. 3; forthcoming; Ripley 2011; Priest 2006, Ch. 20].

In the next couple of sections I will go through some of these extended Liars in detail, as well as the dialetheist's answers to them. There are two types of

[18]Fortunately, nothing here hangs on the understanding of what rationality or rational acceptance amounts to, so I use this simplified intuitive picture without further explanation.

[19]It is generally accepted that we require that one has to have the linguistic means to express these kinds of cognitive states. See [Tappenden 1999] for an alternative view.

the answers given. The first one amounts to saying that the producer of the extended Liar misrepresented or misunderstood the dialetheist's position in some important way. Namely, the definitions of dialetheic semantic terms or principles that lead to the paradoxes are shown to be incorrect, i.e., the object language definitions of *falsity simpliciter* are criticised. The second type of reply dismisses the extended Liar as question-begging, because the theorems and rules used in its formulation are invalid in the given paraconsistent system. For instance, the derivations making use of the *Disjunctive Syllogism* are not damaging for a dialetheist. I will concentrate on the former type of answers, since they are relevant to the *false only* discussion.

Explicit *false only* operators

One could try to define explicit operators in the language that would express the notions *true only* and *false only*. Bremer [2005, p. 50-55], following da Costa's approach [da Costa and Alves 1977], uses this method:

A	TA	ΔA	FA	∇A	$^\circ A$	$^\bullet A$
1	1	1	0	0	1	0
0	0	0	1	1	1	0
0,1	1	0	1	0	0	1

Table 1.1: Truth-value operators

The operators are interdefinable:

$\nabla A =_{Def} \Delta\neg A$

$^\circ$A $=_{Def}$ $\nabla A \vee \Delta A$

$^\bullet$A $=_{Def}$ $\neg {}^\circ A$

And, although he doesn't mention it: $\Delta A =_{Def} TA \wedge \neg FA$ (we can see this from the truth table).

If one wants to express that a sentence A is *just true* it suffices to say ΔA.

The problem is that we can create an extended Liar from any similar formalism. Consider a sentence $A \leftrightarrow \nabla A$[20]. We can't consistently say that A is *false only*[21]:

1) If A is *true only*, then ∇A is *true only*. If ∇A has a value *true*, then A must have value *false only*, contrary to the assumption.

2) If A is *false only*, then ∇A is *false only*. If ∇A has value *false*, then A can be either *true only* or *both true and false* contrary to the assumption.

A will be *false only* and *both true and false*, if we accept a paraconsistent solution to this paradox and say that A is *at least false* or *a dialetheia*. As a result, the distinction between *falsity simpliciter* and *at least falsity* doesn't make a difference.

False only **through valuation functions**

[Smiley 1993; Everett 1993] use the valuation functions of LP's semantics [p. 5 of the present work] to do define *false only* and construct an extended Liar[22].

Define "the value of A" as "$Val(A)$":

1. $Val(A) = \{1\}$ iff $T(A) \wedge \neg F(A)$
2. $Val(A) = \{1, 0\}$ iff $T(A) \wedge F(A)$
3. $Val(A) = \{0\}$ iff $\neg T(A) \wedge F(A)$

4. Define λ as $Val(\lambda) = \{0\}$

[20]For our purposes it is sufficient that *Modus Ponens* holds for an arrow, which is a necessary condition for a "appropriate" conditional for $LP^{\rightarrow}$ discussed in Chapter 2 [See pp. 29-32 of the present work].

[21]See Ripley [2011] for reasons to think that an explicit *just false*-operator inevitably causes trouble.

[22]See [Priest 2006, pp. 287-280] as well.

5. $Val(\lambda) = \{1\} \lor Val(\lambda) = \{1, 0\} \lor Val(\lambda) = \{0\}$ (Every sentence has to be either *true only*, *both true and false* or *false only.*)

6. Case 1: $Val(\lambda) = \{1\}$

7. $Val(\lambda) = \{1\}$ iff $T(\lambda) \land \neg F(\lambda)$ (Instance of 1 for λ)

8. $T(\lambda) \land \neg F(\lambda)$ (6, 7, *Modus Ponens*)

9. $T(\lambda)$ (8, $\land$-Elimination)

10. λ (9, T-scheme)

11. $Val(\lambda) = \{0\}$ (10, Definition of λ)

12. $\{0\} = \{1\}$ (6, 10, Properties of $=$)

13. $0 = 1$ (Extensionality of sets)

14. The same reasoning applies to the second case

15. Consider the third case where Val(λ) = $\{0\}$

17. Val(λ) = $\{0\}$ iff $\neg T(\lambda) \land F(\lambda)$ (Instance of 3 for λ)

18. λ (16, Definition of λ)

19. $T(\lambda)$ (18, T-scheme)

20. $F(\lambda) \lor \neg F(\lambda)$ (Consequence of definitions 1-3)

21. Case 1: $F(\lambda)$

22. $F(\lambda) \land T(\lambda)$ (19, 21, $\land$-Introduction)

23. $Val(A) = \{1, 0\}$ (2, 22 MP; return to 14)

24. Case 2: $\neg F(\lambda)$

25. $\neg F(\lambda) \land T(\lambda)$ (19, 31, $\land$-Introduction)

26. $Val(A) = \{1\}$ (1, 22 MP; return to 6)

27. $1 = 0$ (24, 26, The same reasoning as in 6-15)

Against this derivation Priest doesn't mount any proof-theoretic criticism, but argues that the definitions given in the clauses 1-3 are incorrect, since they overlap. One can't define *falsity simpliciter* in such a way.

> Bearing in mind the original version of the argument, we should expect to have $\mathrm{T}(\lambda) \wedge F(\lambda) \wedge \neg T(\lambda)$, and so both $\mathrm{T}(\lambda) \wedge F(\lambda)$ and $\neg T(\lambda) \wedge F(\lambda)$. That is, cases 2 and 3 of the definition of Val overlap. This is no, therefore, a good definition (any more than is a definition of a numerical function, *f*, such that $f(\mathrm{n}) = 1$ if $\mathrm{n} \leq 5$ and $f(\mathrm{n}) = 0$ if $\mathrm{n} \geq 5$).
>
> [Priest 2006, p. 288]

Further, Priest says that such derivations presuppose the existence or hierarchy of the values, but there are only two of them, true and false, "both" is just an instance, where a sentence gets two values at the same time, it is not a third value. On can talk formally about this intuition using valuation relations instead of valuation functions. Now we have the following definitions:

$Rel(A, 1)$ iff $T(A)$

$Rel(A), 0)$ iff $F(A)$

Define *false only* using valuation relations

[Bromand 2002] came up with the extended Liars for this definition as well, which suggests it can't be used for defining *false only* either. He argues that if the principle (*) *Every sentence is either only true, or only false, or true and false* can be adequately expressed, then the extended liar is inevitable. The most natural and adequate way to express (*) Bromand takes to be[23]:

1. $(v(A, 1) \wedge (\forall x)(v(A, x) \rightarrow x = 1)) \vee ((v(A, 0) \wedge (\forall x)(v(A, x) \rightarrow x = 0)) \vee (v(A, 1) \wedge v(A, 0) \wedge (\forall x)(v(A, x) \rightarrow (x = 1 \vee x = 0))$

Then:

[23] Take $Rel(A, 1)$ to be $v(A, 1)$

2. $v(A, 1) \leftrightarrow A$ (Truth schema)

3. $\lambda \leftrightarrow v(\lambda, 0) \wedge \forall x(v(\lambda, x) \rightarrow x = 0)$

4. $(v(\lambda, 1) \wedge (\forall x)(v(\lambda, x) \rightarrow x = 1)) \vee ((v(\lambda, 0) \wedge (\forall x)(v(\lambda, x) \rightarrow x = 0)) \vee (v(\lambda, 1) \wedge v(\lambda, 0) \wedge (\forall x)(v(\lambda, x) \rightarrow (x = 1 \vee x = 0))$ (1, UI)

5. $v(\lambda, 0) \wedge (\forall x)(v(\lambda, x) \rightarrow x = 0)$ (Assume)

6. $v(\lambda, 1)$ (2, 5, Truth-scheme, MP)

7. $v(\lambda, 1) \rightarrow 1 = 0$ (5, $\wedge$-Elimination, UI)

8. $1 = 0$ (6, 7, MP)

9. $(v(\lambda, 1) \wedge (\forall x)(v(\lambda, x) \rightarrow x = 1))$ (4, Assume)

10. $v(\lambda, 0) \wedge (\forall x(v(\lambda, x) \rightarrow x = 0))$ (3, 9, $\wedge$-Elimination, Truth-scheme, MP)

11. $(\forall x(v(\lambda, x) \rightarrow x = 0))$ (10, $\wedge$-Elimination)

12. $v(\lambda, 1) \rightarrow 1 = 0$ (11, UI)

13. $1 = 0$ (9, 12, $\wedge$-Elimination, MP) (The same reasoning applies to the last case of 4.)[24]

Here Priest gives a reply similar to the one to the extended Liar with valuation functions: the definitions involved are not correct from the dialetheist point of view. (1) doesn't express "the basic semantic principle of dialetheism", as Bromand claims. It is expressed rather by principles 1.a and 2.a:

1.a $(v(A, 0) \wedge \neg v(A, 1)) \vee (v(A, 0) \wedge v(A, 1)) \vee (v(A, 1) \wedge \neg v(A, 0))$

2.a $\forall x(v(A, x)) \rightarrow (x = 1 \vee x = 0))$

The problem is that these principles don't allow to express that something is *false only*, since in this view dialetheias will be *false only* as well.

By 'A is only true', here, he [Bromand] means $(\forall x(v((\lambda), x) \rightarrow$

[24] Here I take Priest's agreement that there are no problems in this derivation as an excuse not to pose question about the restricted quantification for the relevant arrow used [cf. Beall et al. 2006].

> $x = 0$). Quite so. If that is what you mean by 'true only', it does not... Indeed, it is not true of λ by dialetheic lights. Since $v((\lambda), 1)$ is true and $1 = 0$ is not, the conditional $v((\lambda), 1) \to 1 = 0$ and so $(\forall x(v((\lambda), x) \to x = 0)$, is not true. Of course, in a dialetheic context this does not rule out its being true as well, but it does show that there is no legitimate presumption of this[25]
>
> [Priest, p. 294]

If one decides to accept such an answer, she is to be contented with "at least" values. This is technically a less damaging result than extended Liars, since one avoids triviality. But she does so at the similar costs as other solutions to the paradoxes: there are some notions that seem to be expressible (*false only*, *true only*), but are ignored in the theory. Furthermore, these notions seems to be necessary to express genuine disagreement, so one has to look for solutions elsewhere.

In a nutshell, the basic dialetheist response to the extended Liar is to say that such a notion as "false only" doesn't make sense in a paraconsistent object language, nor does Boolean negation, because they lead into trouble in the presence of the "more important" notions (like Truth) [Priest 1990; 2006, Ch. 5; Restall 1999; Beall 2009, Ch. 1 and 5]. Still, if these theorists want to be able to express genuine disagreement (and they usually do – due to its relation to rationality), they are to look for a solution elsewhere. This is the path that Priest [2006; 2008] and Beall [2009; 2013; 2013a] choose to follow: instead of trying to make the changes in the logic and make it tolerate some kind of exclusive connective, they appeal to non-logical (or "extra-logical")

[25] By 'A is only true', here, he means $(\forall x(v((\lambda), x) \to x = 1)$.

theories, such as speech act theory and the theory of rationality. I will turn to these solutions after a discussion of the last "logical" resort these authors point to: formalizing the genuine disagreement using a stronger kind of negation, the "falsum"-negation, which avoids triviality by being defined through the paraconsistent-appropriate relevant conditional. Finally, I will discuss why the dialetheists find this solution unsatisfactory and argue that there might be no *need* to give up on this last resort. The sections to follow will explain why we would *want* to adopt it, especially given the problems arising with the "extra-logical" approaches.

Chapter 2

The *falsum*-negation as a solution to the *false only* problem

2..4 Introduction

In the previous chapter I discussed what would happen if one tried to formalize denial using the exclusive operator "false only" in a paraconsistent logic LP: usually, triviality follows unless we make changes in the logic and give up on some intuitively plausible principles (such as *Modus Ponens* or *Disjunctive Syllogism*, *Modus Tollens*, *Substitution of Identicals*, etc.). But there is still one "logical"[1] possibility that is considered more viable than others, at least by Priest [Priest 2008, Ch. 6; 2006, Ch. 20]: one could use *falsum*-negation in order to express the genuine disagreement, which would play the role of a "false only" operator (i.e., express that some statements are false and not true, and

[1]In contrast to appealing to pragmatics, that is appealing to how an expression can be used as opposed to what is the semantic content of that expression.

consistently so). It means that we add to our language (LP) a constant called *falsum*, "bottom" or **BAD** and written as $\perp$. It stands for a proposition that is always false or for $\forall x T(x)$ ("all sentences x are true"), that is, the trivial situation. Using this formalism we can say that to deny a means not to assert the negation of a, but to assert its *falsum*-negation, i.e., $a \rightarrow \perp$[2] [Priest 2008, pp. 84-85, 104-105].

Priest doesn't take this option too seriously: for him it is "as good"[3] as the classical way of expressing "just falsity" and denial. The criticism of this proposal by Beall [2013a] explains why this is not a very attractive option for a dialetheist. Roughly, in his view, *falsum* is inadequate in expressing what denial really means: *falsum*-negation of a means that *either a is not true, or triviality follows*, which is not what we usually want to express when we deny something.

One might think that such a criticism is the result of the incorrect interpretation of the arrow in the *falsum*-negation as the material conditional ("either-or" clause), which is known to be inappropriate for paraconsistent logic. Of course, Beall has in mind some other kind of conditional when he mounts his criticisms, not the material conditional. But for him this fact makes things even worse – for those conditionals involve a strong logical relation between antecedent and consequent that doesn't seem to be presupposed by most of the acts of denial. Beall doesn't spend much time discussing the *falsum*-negation and its problems, and although intuitively the criticism may seem to be convincing, it is not entirely clear from his discussion what exactly it amounts to when put into the context of the relevant conditionals. In this

[2]We have $\sim \alpha, \alpha \vdash \perp$ and $\perp \vdash \beta$ [Priest 2008, Ch. 4.7; 2006, Ch. 8.5], where $\sim \alpha$ is $\alpha \rightarrow \perp$. In classical logic $\sim$ is equivalent to the Boolean negation.

[3]Which for him might mean "as bad".

section I will analyse this criticism to see what in particular goes wrong with the relevant conditional of the *falsum*-negation.

Here I will consider two types of "arrows": the "strict conditional" of [Priest 2006, Ch. 6] that involves a binary relation on possible worlds, and a ternary-relation conditional used, among others, by Beall [2009] and [Priest 2006, Ch. 20]. While the criticism seems to be fair with regards to the former, the latter looks more promising, as it obviously helps to avoid at least a part of the problem. The worry here is that it might create other problems – for there might be difficulties involved in interpreting the relevant conditional in an intuitive way [cf. Copeland 1979]. I will turn to [Beall et al. 2012] for the framework that allows to give the intuitive picture of the ternary relation conditional and use it to address the above criticism that the *falsum*-negation is too strong an operator to adequately express denial.

2..5 Priest on *falsum*-negation

First, I will talk about the peculiar place the discussion of the *falsum*-negation has in Priest's work: its role as a defence against the "classicist's" criticism explains weakness of Priest's claims. Then, I will introduce Beall's criticism.

Priest argues that the *falsum*-negation is a way of expressing denial that is not worse (but maybe not better either) than classical logic negation [Priest 2008, Ch. 6; cf. Priest 1990]. The particularity of this argument consists in the fact that he considers the *falsum*-negation only as an answer to the "classical" criticism that might be summarized as follows: "the paraconsistent logic is worse than the classical logic, because it lacks the means of expressing genuine disagreement (i.e., that some claims are "false only") the classical logic has (using Boolean negation)". Here the discussion of the *falsum*-negation is a defensive manoeuvre, which explains why he is happy with the (rather weak, considering his attitude towards the classical logic) conclusion that the classical logician is no better than him with respect to expressing denial. The minor discursive role of this argument explains why Priest doesn't spend a lot of time exploring this possibility, nor debating the criticisms of his critics on this point.

> A dialetheist can express the claim that something, A, is not true in those very words, $\neg T\langle A\rangle$. What she cannot do is ensure that the words she utters behave consistently: even if $T\langle A\rangle$ holds, $\neg T\langle A\rangle$ may yet hold. But in fact, a classical logician can do no better. He can endorse $\neg T\langle A\rangle$ but this does not prevent his endorsing A as well... Classical logic, as such, is no guard against this... All the classical logician can do by way of saying something to indicate that A is not to be accepted is to assert something that $\neg A$ will collapse things into triviality if he does accept A. But the dialetheist can do this too. She can assert $A \rightarrow\perp$.
>
> [Priest 2006, p. 291]

The dialetheist is thus not in a really worse position than the "classicist" with respect to the "false only" problem. The classical logician only seems to fare better and be able to say that some statements are false *simpliciter* in a natural way. In fact, both have to do the same thing in order to express their disagreement: they have to appeal to the trivial models where *falsum* holds. The difference is that the paraconsistent logician has to do so explicitly by asserting $A \rightarrow\perp$ when she thinks that A is *just false*, whereas the classical logician can convey the identical message by asserting $\neg A$.

Although there is this similarity between the paraconsistent and classical approaches, it is worth mentioning a couple of important differences. First, in contrast to the dialetheist, for the classical logician, the notions of negation, *falsum*-negation, and denial are equivalent. In the classical framework one can express denial of A by saying "not A", which is the same as stating that *either A is false, or triviality follows*. For the dialetheist[4] these meanings are not equivalent: *falsum*-negation is stronger than the usual, "at least not", De Morgan negation, and only the former might be expressing the act of denial, a signal that one's beliefs have to be revised if the denied statement is already present in her belief set. In the dialetheic theory, the meaning of negation and *falsum*-negation bifurcate due to the fact that some rules governing the *falsum*-negation are dropped. For instance, the *Law of Excluded Middle* is not supposed to hold for the *falsum*-negation on pain of triviality (a Strengthened Liar can be formulated otherwise). *LEM*'s failure fits well the *falsum*-negation's purpose of expressing denial, as the latter seems to allow for the gaps: one might withhold an opinion on some subject, i.e., neither deny nor assert some proposition (without the irrational commitment to deny and assert it at the same time)[5]. Second, the conditional used in formulating the *falsum*-negation is not the material conditional, again, on pain of triviality (given the Curry paradox[6]). These restrictions are supposed to make the *falsum*-negation, in contrast to the Boolean negation, a meaningful notion, one that doesn't trivialize the language and, moreover, fares well (as well as

[4]It is worth noting that it is true not only of the paraconsistent logician, but the paracomplete ("gappy") theorist as well.

[5]I will talk about the exclusiveness of denial and assertion and its relation to rationality in the last section of this chapter.

[6]Here the Liar is just an instance of the Curry paradox: $T\langle A\rangle \leftrightarrow (T\langle A\rangle \rightarrow\bot)$.

the classical negation) with respect to formalizing denial. This conclusion is supposed to tip the scales in favour of the paraconsistent logic.

JC Beall grants the first part of this conclusion, but questions the second one[7]. The reason for that is the conditional used in $A \rightarrow \perp$. In his view, the logical connection between A and $\perp$ expressed by the relevant arrow is much stronger than one needs in order to deny A. We read in [Beall 2013a]:

> Priest's reply to the just-false problem involves a logical-strength connection between the "just false" antecedent and the explosive consequent. The problem is that $A \rightarrow \perp$ is virtually never true! On Priest's given account, with the corresponding semantics, the glut theorist would truly say that A is "just false" exactly when there is no non-trivial world whatsoever, no (non-trivial) possible or impossible world in which A is true. That would do the trick of indicating that you take A to be acceptable on pain of accepting the trivial theory; but it's too much – much too much.
>
> [Beall 2013a, p. 4]

The clause "A is "just false" exactly when there is no non-trivial world whatsoever, no (non-trivial) possible or impossible world in which A is true" corresponds to the understanding of conditional presented by Priest in the first [1987] edition of [Priest 2006, Ch. 6]. This is a strict conditional, defined in terms of a binary accessibility relation on possible worlds. Since the kind of arrow to replace the material conditional in LP is not a settled question, one would make more out of the Priest's argument by looking at what would hap-

[7] He might agree that it does in fact fares *as well* as the classical negation, but, from his point of view, the latter fares not well at all.

pen if we interpreted "$\rightarrow$" in a more "modern" way[8]. Moreover, Priest himself makes some changes to the discussion of entailment in the second edition. There he takes an "appropriate conditional" to be represented by the relevant arrow based on the ternary accessibility relation on possible and impossible worlds. [ibid. Ch. 19.8].

First, let's have a look at the conditional appealed to in the quote[9]. In a nutshell, this is the conditional that fixes the problems with the material conditional that arise when it is put in a paraconsistent setting (LP). Among those are the failure of *Modus Ponens* and *Substitution of Equivalents*, as well as the possibility of formulating the Curry paradox[10]. The common paraconsistent way of achieving these goals is by appealing to the binary or ternary relation on possible, and sometimes impossible, worlds (points, situations, information states, etc.) when giving the truth-conditions for the conditional ("Routley-Meyer-Priest approach").

Priest [2006, Ch. 6.3] gives the following truth-conditions for the conditional. A semantic interpretation for the language of LP is a quadruple $M = \langle W, R, @, v\rangle$, where W is a set of possible worlds, R is a binary relation on W, @ is a special member of w, the "actual world", and v is an evaluation of the propositional parameters, i.e., a map from $W \times P$ (P is the set of propositional parameters) into $\mathcal{V} = \{\{1\}, \{0\}, \{1, 0\}\}$[11]. Then:

[8]Here I will look at Beall's [2009] presentation of such a conditional. This is motivated by the fact that he is using the same approach as Priest does: he takes LP as a basis and then adds a relevant arrow to it. This makes the comparison easier.

[9]I assume that it is the conditional of the chapter 6 [Priest 2006]. There is no appeal to the impossible worlds there, but one could easily divide the domain of worlds into these two categories and define their validity separately (see below). But given that here one uses the strict conditional based on a binary accessibility relation, such a distinction would be redundant.

[10]See 4.9 of [Priest 2006] for the reasons why the conditional should be non-contraposible as well.

[11]The clauses for other connectives are reformulated in terms of possible worlds as well in

$1 \in v_w(A \to B)$ iff for all w' such that Rww', if $1 \in v_{w'}(A)$ then $1 \in v_{w'}(B)$, and if $0 \in v_{w'}(B)$ then $0 \in v_{w'}(A)$.

$0 \in v_w(A \to B)$ iff for some w' such that Rww', $1 \in v_{w'}(A)$ and $0 \in v_{w'}(B)$.

One would also want to add some special constraints on R in order to make this definition suitable as an interpretation of the "appropriate conditional" that validates the intuitively correct principles. For instance, Priest requires the actual world @ to be "omniscient", i.e., that for all $w \in W$, $R@w$. This constraint on R allows us to validate *Modus Ponens*: it holds just in case R is reflexive and, *a fortiori*, given the "omniscience" of @[12].

Given above conditions, denial as the *falsum*-negation will be interpreted in the following way:

We are to deny A in @ iff $A \to \perp$ is true in @, i.e., iff for all w' such that $R@w'$, if $1 \in v_{w'}(A)$ then $1 \in v_{w'}(\perp)$, and if $0 \in v_{w'}(\perp)$ then $0 \in v_{w'}(A)$.

This means that in all accessible worlds where A is true, $\perp$ is true as well, which in turn means that A is to be denied just in case when it can't be true in any accessible world except for the trivial one, where everything is true[13].

Now we can see what drives the criticism of the *falsum*-negation approach to the "just false" problem: such a picture requires too strong of a condition on denial. We can deny that the grass is fluorescent blue, but still acknowledge

a straightforward way. For instance, $1 \in v_w(A \wedge B)$ iff $1 \in v_w(A)$ and $1 \in v_w(B)$. Appeal to worlds, w, is redundant here, as the truth-conditions don't depend on the possible worlds other than w. See [Restall 1999] for definitions of De Morgan negation using the Routley-star semantics.

[12]See [Priest 2006, p. 88] for the list of principles validated in the presence of this constraint, among them are *Modus Ponens* in the rule form, *Suffixing* and *Prefixing*, *Reductio*

[13]If one takes $\perp$ as always false, then it has to be impossible for A to be true *tout court*.

the fact that it is possible that the grass could look this way (unless one puts unlikely strong constraints on what is possible). Given this definition of arrow, the criticism seems to be well-motivated: under this interpretation, although sufficient for a denial of a statement, the *falsum*-negation is definitely not necessary.

2..6 Ternary relation conditional, the first attempt

What happens if we interpret the conditional in a different way? An arrow based on the ternary relation on possible (and impossible) worlds seems to solve the above problem for the "Priest's conditional" that the *falsum*-negation is virtually never true, i.e., it is true only in the trivial worlds. [Beall 2009, Ch. 2] gives a short presentation of one such arrow:

> The idea, then, is that the normal condition... governs our conditional at all normal... points: $A \Rightarrow B$ is true at a normal world iff there is no world... at which A is true, but B is not. At abnormal points, the conditional is constrained by the given ternary relation: $A \Rightarrow B$ is true at an abnormal point w iff there is no w-accessible pair $\langle w', w'' \rangle$, such that A is true at w' and B not true at w''. Whereas the normal condition (now restricted to normal worlds) involves looking only at (all) points taken by themselves (and checking whether the consequent is true at the point if antecedent is true), the abnormal condition involves looking at *pairs* $\langle y, z \rangle$ of points and checking whether the consequent is true at z if antecedent is true at y. Of course, sometimes, $y = z$, in which case one is back to checking a point "by itself", but sometimes $y \neq z$.

[Beall 2009, pp. 29-30]

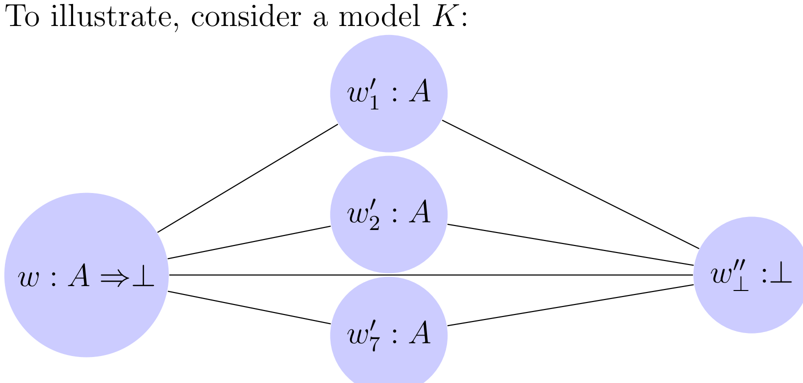
Now we have to add another member N (the set of "normal worlds") to our interpretation M, $M = \langle W, N, R, @, v\rangle$. Then the clauses for the conditional will be the following[14]:

> Where $w \in N$ and $w' \in W$, $1 \in v_w(A \Rightarrow B)$ iff for all $w' \in W$, such that Rww', $1 \in v_{w'}(B)$ if $1 \in v_{w'}(A)$;
>
> Where $w \in W - N$ and $w', w'' \in W$, $1 \in v_w(A \Rightarrow B)$ iff for all $w' \in W$, such that $Rw\langle w', w''\rangle$, $1 \in v_{w''}(B)$ if $1 \in v_{w'}(A)$.

In comparison to Priest's arrow, things are different for the *falsum*-negation only in the "abnormal" $(W - N)$ case. Denial will have the following meaning given such truth-conditions for the arrow:

> If we are to deny A in an "abnormal" world[15] w, then we are to assert $(A \Rightarrow\perp)$ at w, which is true iff for all pairs of worlds accessible from w ($w', w'' \in W$, such that $Rw\langle w', w''\rangle$), $\perp$ is true at w'' if A is true w'.

To illustrate, consider a model K:



[14] I change the notation to be closer to Priest's.

[15] Some suggest that @, our actual world, might be one of the abnormal ones [Priest 2006, Ch. 20]. One can do without the possible/impossible worlds distinction and use the ternary relation for all "worlds" (make constrains on the models) [Mares 2004, pp. 210-211 on Routley and Meyer 1973]. Nothing hangs here on this.

We have five worlds, A is true at three of them w'_1, w'_2, w'_7, $\perp$ is true at $w''_\perp$, our designated trivial world, and we want to deny A, i.e., assert $A \rightarrow\perp$ at "our point" w. It is easy to check that we can correctly do so given the above truth-conditions.

This interpretation of the conditional might allow for a more plausible interpretation of denial, as it doesn't require A to be true only at the trivial worlds (worlds, where $\perp$ is true[16]). In K, for instance, we see that the worlds where A is true can be either trivial or not (A is true at the non-trivial w'_2 and trivial $w''_\perp$). This solves the problem that we had with binary relation account: we can deny that the grass is fluorescent blue in the world w, but imagine (or "access") the possible non-trivial worlds that have fluorescent blue fields. Here the requirement is weaker: the A-worlds[17] has to be related to triviality in some way (indirectly through w or directly). The question is how to interpret the relation between the $A \rightarrow\perp$-world, A-world and the $\perp$-world[18], so that it makes sense as the explanation of the meaning of denial.

Possible world semantics don't seem to be very promising if we decide to treat them as more than just formal tool. Saying that you can deny A just in case when all possible worlds where A is true are related only to the trivial worlds doesn't clarify much. You will have to interpret this peculiarity of A-worlds in some way. Why should a possible world that differs from ours by the fact the grass there is typically fluorescent blue have to be related to the trivial world? There doesn't seem to be an intuitive answer to this question[19].

[16]Here I will assume that there is only one such world, for the sake of simplicity.

[17]An "A-world" means a world where A is true.

[18]Alternatively, relation of w to the pairs $\langle w'_1, w''_\perp\rangle$, $\langle w'_2, w''_\perp\rangle$, $\langle w'_7, w''_\perp\rangle$, etc.

[19]If we treat the possible worlds as arranged by some kind of "alikeness" relation, it seems to make sense to say that the A-worlds of the model K are the least alike to the actual world w, that they are the "oddest" worlds, since they are related to the $\perp$-world (the latter being the extreme of the "alikeness" spectrum), which explains why A should be denied. But this

Alternatively, we could say that even though there are possible worlds w'_1, w'_2, w'_7, where it is reasonable to assert A, asserting it at w would lead to triviality, simply because *Modus Ponens* is valid (because we would have $A \Rightarrow\perp$ true at w). Denying A at an A-world would have the same effect for the same reason: given the reflexivity of R (that validates MP), we have that the A-world is related only to the $\perp$-world and an A-world. This helps avoiding the odd metaphysical assumptions, but, being entirely formal, it doesn't explain the nature of the relation between A and $\perp$. Moreover, MP usually presupposes a causal link between the antecedent and consequent, which is much stronger than we would want for denial.

We are in need for a weaker and more plausible understanding of the relations between the "worlds". At this point it can be helpful to look at some other interpretations of the ternary relation – they might shed light on how to interpret denial in this framework in a intuitive way.

---

won't work. For instance, think of a world w' with the blue grass, where everything else is like in our world. We would want to deny that the grass is blue at our world w, but it is implausible that there are no worlds odder than w', that it is close to the trivial world. Here David Lewis could say that there couldn't be a world "just like ours except for the florescent blue grass". He would claim that lots of other things would have to be changed also (including some scientific laws). It is fair to suppose that there might be interconnected facts that cannot be changed but together: for example, a world where the only change is that the average temperature is $4°C$ higher than now will look very different from the actual world. Here, denying that we are living in the world where the average surface temperature is $18°C$ could be plausibly interpreted as asserting that the $18°C$-world is not at all "alike" to ours, that it is too "odd". But this is not true of every fact. We lived in a $14°C$-world not so long time ago and one can't claim that this world is as odd as $18°C$-world, although the fact that we now live in a $14°C$-world is correctly deniable. On this picture, one would have to prove that all facts have very strong "butterfly effects" in order to relate denial and triviality. This is the kind of metaphysical claims I am trying to avoid when using information states interpretation.

Maybe people who are more familiar with the philosophical tradition of interpretation of the possible worlds semantics might come up with a plausible interpretation of this framework; in this case, the following discussion of the alternative "information links" interpretation may be considered as an invitation to compare the two.

2..7 The intuitive interpretation of the ternary relation conditional for the *falsum*-negation

Here I will consider one of the common interpretations of conditional analyzed in [Beall et al. 2012], namely modal (absence-of-counter-example) conditional, which is based on the notions of information and information links, as well as the general framework of relative relative possibility[20].

According to this interpretation, take it that our ternary relation $Rww'w''$ holds just when w'' is possible relative to w', relative to w.

A point w is possible relative to w' iff everything required (necessary) at w holds at w', i.e., iff for every constraint ("A requires B") whose antecedent holds at w, its consequent holds at w'[21]. And whether w'' is possible relative to w' can only be answered relative to a point w with its "constraints". In other words, w "hosts" the constraints – inferential relations between the propositions – and w' and w'' "realize" those constraints or "links" [ibid., pp. 12-13].

One of the most common understanding of a conditional appeals to the absence of the counter-examples. Namely $A \to B$ holds when there is no situation (possibility, point, world, etc.) where A is true and B is false, i.e when there is no counterexample to the given inference. The ternary relation framework interpreted through the relative relative possibility easily accommodates this common definition:

[20] They survey three common interpretations, I skip two ("conditional logics" and "conditionals as operators") as they are either not relevant or redundant for this discussion.

[21] "Constraints" are the conditionals that hold in the actual situation or "world", such as scientific laws, rules, conventions. One can think of w as a Turing machine head, where we have the transition rules, w' as the input and w'' as the output – the result of application of rules from w to w'. In other words, w'' shows the outcome required by input w' given the rules w.

> For the pair point $\langle yz \rangle$ to count as a counterexample, according to x, is simply for z to be relatively possible from y, according to x: if everything required at y, according to x, holds at z, then if A holds at y but B doesn't hold at z, A can't require B according to x.
>
> [Beall et al. 2012, p. 13]

This framework does so by appealing to the unfamiliar "pair points", though. The theory of "situated inference" [cf. Mares 2004; Restall 1999] allows for an intuitive explanation of what the pair points and their relation to the "usual points" are: the pair points are the objects that realize the "information links" (think of the laws of nature, conventions, rules, etc.). Here the "worlds" or "points" are understood as the information states that provide you with the pieces information (A, B, C), but also "information links" (what follows from what, i.e., $A \rightarrow B$, etc.). When we apply the analysis of conditional in such states to the analysis of denial as a *falsum*-negation a[22], we get:

> Suppose we are in a situation w in a world W (the set of situations), which contains an informational link that *a carries the information* $\perp$. Then from the hypothesis that there is a situation w' in W in which A is true, we can infer that there is a situation w'' in W in which $\perp$ is true.

In terms of absence-of-counterexample:

[22] "The information in a situation not only tells us about that situation; it allows us to make inferences about other situations in the same world. Suppose that we are in a situation x in a world w and that x contains an informational link that says that A *carries the information* B. Then from the hypothesis that there is a situation y in w in which A is true, we can infer that there is a situation z in w in which B is true" [ibid. p. 7].

> There is a counterexample to $A \to B$ in w iff there is a point pair $\langle w', w'' \rangle$ that realizes the informational links of w, but $1 \in v_{w'}(A)$ and $0 \in v_{w''}(B)$. Saying that $A \to \perp$ holds at w would mean that there is no point pair $\langle w', w'' \rangle$ that realizes the informational links of w, but $1 \in v_{w'}(A)$ and $0 \in v_{w''}(\perp)$.

The problem is that it is not plausible that any statement "carries" the trivial information. If "carrying" presupposes containing the information being carried, as it seems to do, it is difficult to see what sentence would "carry" the triviality other than $\perp$ itself – it is "much – too much" for the denial. We are in need of a weaker understanding of "informational link" that ties together a denied A and the $\perp$ than the "carrying of information". The relative relative possibility framework is helpful in this respect.

Here, for the pair point $\langle w', w'' \rangle$ to count as a counterexample, according to w (according to the available "informational links"), is for w'' to be relatively possible from w', according to w. If everything required at w', according to w, holds at w'', then if A holds at w' but B doesn't hold at w'', A can't require B according to w. $A \to B$ would be true if there is no pair point counterexample to it[23]. Does this definition help to make sense of the denial as *falsum*-negation?

The basic idea behind the ternary relation $Rww'w''$ in this framework can be reformulated in the following way: if we were to complete a situation w' according to (the information links, rules, etc.) w, then we would get the

[23]For an example think of Peano Arithmetic axioms as our situation w, where we want to check if some conditional statement $4 < 5 \to 4 + z < 5 + z$ is true. We know that $x < y \to x + z < y + z$ holds at w, and w'' has the consequences, according to PA axioms (w), the fact $4 < 5$ would have if it were true (at w'), such as $4 \not> 5$, $5 > 0$, etc. If among these consequences at w'' we don't find a counter-example to the consequent (such as $4 + 4 > 5 + 4$), our conditional is true.

situation w''. The counter-factual aspect of such a definition ("if we were to...") makes it weak enough for our purposes. In our simple example K, when one is willing to deny A at w (i.e., assert $A \rightarrow\perp$ at w) she intends to convey that there is an information link available that "says": according to our situation w, if we were to assert A in the possible situation w' and "realize" (assert) everything required by this new situation, then we would find ourselves in the trivial state w''. Although this reformulation seems to be quite similar the previous one (in terms of "carrying the information"), it gives more flexibility to the "link" that ties together the antecedent and consequent of the *falsum*-negation of A.

In order to make it suitable for the explication of denial we can understand the "realize everything that is required by w" as meaning to construct a maximal set of propositions w'' from w'. We do it by adding all the propositions that could be denied and asserted given the "information" or "possibilities" of w', as well as the "information links" and "rules" given in w. The idea is that we "complete" w' or make explicit what is implicitly accepted at w' in this special set or situation w''. On this picture even though it is possible that there are situations where A is acceptable (assertible) without leading to triviality, if you assert it from the perspective of the point where it is denied (w), then – by making everything explicit that was implicit in the situation – you could get to a situation where anything would be assertible (there is a situation that realizes or follows through everything required by this contradiction and it is trivial).

Now one is, of course, to explain why everything follows from this contradiction, from the fact that one denies and accepts A. It might seem surprising and unilluminating – haven't we just re-introduced explosion trying to avoid it (In other words – isn't this definition is too strong again?).

In order to address this concern one should look at the dialetheic understanding of assertion and denial as the linguistic expression of the cognitive acts of acceptance and rejection, and their relation to belief revision. According to the classical picture one is to revise her beliefs if there is a contradiction in the belief set (A and its negation), while for the dialetheist, who can tolerate the statements of the form $A \wedge \neg A$ in her belief set, this role is played by the mutually exclusive pair of assertion and denial. It means that if one finds herself in the situation where she is denying and asserting the same proposition, she has to revise her beliefs and either withhold her belief with respect to the proposition in question or accept the option that is better supported (by evidence, accordance to the cognitive values – whatever one takes to be decisive for the rational acceptance). Such role of acceptance and rejection (as well as their relation to the denial and assertion) explains why asserting and denying something at the same time leads to "explosion" of the belief set: being the constraint on the belief revision similar to the principle of non-contradiction for the classical belief revision, "breaking" it cripples the whole system. I will discuss the dialetheic view of rational acceptance, belief revision and its relation to denial and assertion in more detail in the following section.

One might also wonder if this explanation of the arrow in the *falsum*-negation is not *ad hoc*, i.e., applicable only for our purposes. In order to avoid *ad-hocness* in this situation one has to check if such an explanation of the conditional suits the definition of the conditional in general and can be used in *many* situations, not only when it is useful for this discussion.

The idea of possible (or counter-factual) "completion" of the given situation, set or point is quite natural to the intuitive understanding of conditional.

The plausibility of the (relative relative) possibility aspect of the conditionality is already explained by [Beall et al. 2012] (see pp. 12-14), so I will turn to the counter-factual twist I added in order to analyse the denial.

Think of a random conditional "If the grass is fluorescent blue, then it is not fuelling by photosynthesis" that is true for w (say, it is a scientific discovery). By our conditions, we complete a possible situation w' where the grass is really fluorescent blue to a "maximal set" w'' that realizes everything required by w' (all consequences of propositions at w'), according to our discovered "laws". Given that we have the "according to w" clause, the conditional is true and avoids the the extreme counter-factuality (if the laws were different): given the laws of nature of our world, if we were to plant the fluorescent blue grass, it couldn't survive using the photosynthesis. In the situation w' we have the explicit piece of information A, that the grass is fluorescent blue. If we were to "make everything out" of this situation according to the rules of our world w, we would contemplate the situation w'', where among other things (maybe, this grass is not attractive as food for cows and rabbits, or you could make some different chemical use of it – depends on what "constraints" hold at w) this grass is not showing the signs of fuelling by photosynthesis. Basically, every time we check the truth of a conditional we consider this maximal set of consequences of the antecedent – everything that might follow from it given what we know – and see if the consequent is present in this set.

2..8 "Explosion" of denial and assertion

In the previous section I mentioned that "explosion" of denial as the *falsum*-negation has an intuitive explanation in the dialetheic framework, contrary to the expectations one might have about dialetheism. This point was important in supporting the claim that the *falsum*-negation adequately expresses denial (i.e., that the arrow used in $A \to\perp$ is not "too strong"). Here I want to elaborate on this point, as it may appear to be quite surprising within the dialetheic framework and requires further explanation.

The explanation, according to Priest, goes as follows: The pair of assertion and denial has to be exclusive[24], because it expresses what is rationally accepted and rejected; and acceptance and rejection are exclusive. This is considered to be a basic fact about rationality – there is no example when one (explicitly) rejects and accepts something at the same time (it is more convincing under the assumption that different behaviour corresponds to acceptance and rejection of a given proposition, as it is indeed impossible to behave differently at the same moment).

> ...acceptance and rejection do appear to be incompatible. One can certainly believe something and believe its negation. One might even argue that one can believe something and not believe it, though this is much more dubious. But it seems difficult to argue that one might both believe something and *refuse* to believe it. Characteristically, the behaviour patterns that go with doing X and refusing to do X cannot be displayed simultaneously.
>
> [Priest 2008, pp. 98-99]

Moreover, the pair assertion and denial *has to* be exclusive, because of its role in the belief revision process. Without this constraint the "explosion" of the belief set would follow, i.e., one would be justified to accept anything.

[24] I.e., one is not allowed to assert and deny A simultaneously; asserting A implies not denying A, and denying A implies not asserting A.

This view is quite similar to the view of the classically-minded theoretician, for whom the role of such a constraint is played by consistency of the set of accepted propositions. The dialetheist wants to keep the simple classical belief revision picture (an unacceptable contradiction $\Rightarrow$ belief revision), since it ensures that the "genuine disagreement" can happen and be expressed. In order to do so she has to allow for denial and assertion to be exclusive, but in a way that wouldn't prevent adding some true contradictions to the belief set. First of all, this requires some changes in the way we understand the link between negation and denial.

One of the main features of the dialetheic understanding of denial is that denial is not equivalent to the assertion of a negated sentence. This distinction is motivated by the dialetheic theory, as well some independent examples:

> ...the intuitionist who rejects an instance of the law of excluded middle, $A \vee \neg A$, does not, most emphatically, accept its negation, which implies $\neg A \wedge \neg\neg A$. Conversely, one may accept $\neg A$ while failing to reject A. One would do this if, while being convinced that $\neg A$ is true, one acknowledged the possibility that it might be a dialetheia.
>
> [Priest 2008, p. 99]

The desire to separate denial and assertion of negation may also be motivated by the distinction between metalinguistic and content negation [Horn 1989], which makes us wonder if this is the distinction dialetheist is interested in. Metalinguistic negation is used in order to negate the way the content of the sentence was expressed, rather than the content itself. A typical example is the case of presupposition failure: "No, she hasn't stopped cheating (she never started doing it in the first place)". Here, we don't want to assert the contents of negation of the statement "She stopped cheating"; rather, we want to say that "something went wrong" in this sentence. Similarly, Tappenden [1999] uses distinct examples of metalinguistic negation to prove that in natural language denial of a statement is not equivalent to assertion of negation of a given sentence ("Slick Willie did not speak to us. President Bill Clinton did"; "Some men are not chauvinists. All of them are"; "John isn't wily or crazy. He's wily and crazy"). In these examples one doesn't want to assert both propositions that form a contradiction, but rather cancel, correct the presupposition or implicature contained in the first sentence. This might make one think that the cases of metalinguistic negation are the cases of "corrective disagreement". Given that they are similar to what I call corrective disagreement for dialetheist ("The Liar sentence is not true, it is both true and false"), one might wonder if dialetheist's negation in this case is metalinguistic negation.

In order to understand how the distinction between metalinguistic and content negation differs from the types of negation used by dialetheist, let's look at another example of a dialetheia. By Priest's view, when you are in the doorway you are both in the room and not in the room at the same time [Priest 2006, Ch. 12][25]. When Priest wants to correct somebody who mistakenly thinks that he is both in the room and not in the room (and he is definitely outside the room), he *denies* that he is both in the room and not in the room. Here denial is not equivalent to assertion of negation of the statement (as not both true and false equals to both true and false). The difference from the above cases is that here dialetheist expresses genuine disagreement and definitely doesn't use metalinguistic negation (nothing "goes wrong").

[25] I give the detailed explanation of this example on page 58-60 of the present work.

When Priest wants to correct a student who thinks that he is in the room when he is in the doorway, he asserts: "I am not in the room, I am both in the room and not in the room", which looks similar to the "metalinguistic" case. The appearance is misleading though: here assertion of negation doesn't mean denial, but the second part of the sentence adds some information to the first, rather than cancelling it. This is the cases I would describe as the ones involving "corrective disagreement". They differ from Tappenden's metalinguistic examples, since the latter should be analyzed as the cases of denial, not negation ("genuine disagreement"). In fact, for Priest, the sentences where "something goes wrong" are "just false" ("genuine disagreement").

> Suppose that A is a sentence, and suppose that there is nothing in the world in virtue of which A is true – no fact, no proof, no experimental test. Then this is the Fact in virtue of which not-A is true... in the case of denotation failure, we might distinguish between the case where "John's brother is a butcher" is false because John has no brother, and that where it is false because he has a brother who is a French-polisher... if A is any atomic sentence of a kind whose members have been proposed as truth valueless, not-A is true. Thus, "Julius Caesar is not a prime number"... and so on are *simply true*... let us return to the example... "This sentence is true". We saw there that the truth conditions of this sentence imply neither the truth of this sentence nor its falsity. There is therefore no question of an a priori proof (or refutation) of it... Hence, by the previous discussion, this sentence is *simply false* and its negation is true.

[Priest 2006, pp. 65-66]

Within the proposed framework, analysis of presupposition failure (metalinguistic negation) is done by means of denial – we have to deny that Julius Caesar is a prime number. There are some other cases where metalinguistic negation isn't involved, but they are analyzed as denial as well (e.g., the case where Priest denies that he is both in the room and not in the room). It means that denial is not equivalent to what is thought to be metalinguistic negation, neither it is appropriate as an analysis of "corrective disagreement", as I understand it here, since the latter involves assertion of both "denied" statement and its correction[26].

The distinction between denial and assertion is needed to keep the usual link between the truth, assertion and rational acceptance. We want to say that one is supposed to assert what one accepts as true, and that it is rational to do so. But if we accept such a link, as well as understanding of denial as definable by assertion of negation, then the contradiction coming from the Liar paradox becomes a real problem for a rational agent. Think of a Liar sentence L: "This sentence is not true". Given the above commitments, if it is true, we should accept it and therefore assert it. But if it is true then what it says is true as well, so it is not true, which means (by equivalence of assertion of negation and denial) that we are to deny it. The same conclusion follows from supposing that it is not true and that we have to deny it. We end up in a situation where we are supposed to assert and deny the same sentence. But the link between asserting and denying and the acts of rational acceptance and rejection make such a conclusion unacceptable: we are supposed to accept

[26] The sentence "Some men are not chauvinists. All of them are" doesn't assert a contradiction, whereas "The Liar sentence is not true, it is both true and false" does.

and reject something at the same time.

Why is such a conclusion unacceptable for a dialetheist? Why not accepting another contradiction? It seems that if one accepts that A and $\neg A$ can be both true one could allow for denying and asserting both sentences at the same time.

It does seem blatantly impossible to rationally accept and reject something at the same time. But where does this intuition come from? There are answers available if one is willing to commit herself to some theory of cognition (e.g., as Priest points to dependence of actions on cognitive states of acceptance and rejection, and then appeals to the impossibility of acting in opposite ways at the same time). This kind of answer points to the correct direction (showing that the impossibility of accepting and rejecting hangs on some obvious impossibility), but I am not willing to defend either of those theories and I think there is an answer available independently of them. On this view, the impossibility of simultaneously accepting and rejecting a proposition is explained by appealing to the way a rational agent revises her beliefs in the face of new evidence.

One of the most common reactions you can get when somebody says that there are true contradictions is that it is irrational to hold such a view. This is because consistency is usually considered as the foundation of rationality, not only a principle of logic [see the essays in Beall, Armour-Garb, Priest 2004]. Generally, in such a framework [see Priest on belief revision in 2006 for references and formal exposition of the view] beliefs of a rational agent consist of a set of propositions (that may be incomplete) closed under some consequent relation. When the agent comes to acquire new information from a reliable source she adds it to her belief set, there are two possibilities depending of how this belief interacts with the given set. If the belief set doesn't contain the contradictory of the new belief B, the latter can be added to the set. But if we find that the addition results in the set containing B and $\neg B$, the agent has to revise her beliefs, which means that she is supposed to dispose of the old belief or just suspend her belief if the degree of belief for $\neg B$ is high enough. On this classical picture ("the AGM conditions" [Priest 2008, Ch. 8.2; cf. Gördenfors 1988]) accepting both the contradictories means that the belief revision never has to happen, one can just expand the belief set up to triviality. Here, the contradiction "explodes" as it does in classical logic and the explanation for it is that consistency is the constrain on the belief revision that prevents the belief set of a rational agent from being potentially trivial. If one presupposes such a picture of rational belief, then of course true contradictions are irrational (the same holds for classical logic). [See Priest 2006, pp. 103-104; he makes reference to Popper 1940: "What is Dialectic?", pp. 316-317 and Lewis 1982 "Logic for Equivocators"]

Priest has an easy answer to a classical logician who claims that true contradictions are rationally unacceptable, because they lead to triviality. This classical objection is criticized as question-begging, since it presupposes the classical logic as a correct one, where *Explosion* and *Disjunctive Syllogism* are unrestrictedly valid, but it is not the case for many paraconsistent logics, so one has to give some independent reasons to think that either the classical logic is correct or that the given principles are correct. Does Priest have a similarly easy answer to the belief-revision version of the argument?

It seems that he does [see Priest 2006, pp. 103-107]. In this case he says that the above picture of rational acceptance doesn't presuppose the incorrect rules (that inconsistency leads to triviality), but rather incorrect analysis of the relation between acceptance, rejection and negation of the propositional content of those linguistic and cognitive acts. Priest agrees that it is irrational to accept and reject some A at the same time, but this doesn't mean that one can't accept both A and $\neg A$ without rejecting either of them. On this picture rejection is not the same as acceptance of negation of a statement.

This allows to keep the familiar picture of belief revision, where consistency is the main constraint. However, here it is not the consistency of the content of the beliefs, but rather the consistency of the attitudes towards this content, which allows to incorporate the cases where we have contradictory content that we are rationally forced to accept (as in the Liar case). So, if we accept B, i.e., it is in our belief set, when the agent comes to know $\neg B$, then she directly accepts it and adds it to her belief set, but only if $\neg B$ is not already rejected. If it is, then she must revise her beliefs: either to accept $\neg B$ and discard the rejection of $\neg B$ or suspend her judgement about either acceptance or rejection of $\neg B$ if the evidence doesn't determine the decision. Here, acceptance of B doesn't depend on acceptance or rejection of $\neg B$[27].

> Rejecting something...is putting a bar on accepting it [something] (although, of course, one can change one's mind about this in the light of new evidence, etc.). When justified, it is so because there is evidence against the claim: positive grounds for keeping it out of one's beliefs – rather than the mere absence of grounds for having it in.
>
> [Priest 2008, p. 103]

[27] As acceptance and rejection are linguistically expressed by assertion and denial, the possibility of formulating the Liar will depend on the features of the latter. If denial is formalized as *falsum*-negation the paradoxical reasoning doesn't go through, since *Contraction* is not valid in $LP^{\rightarrow}$ and the Liar is equivalent to the Curry paradox. Priest's pragmatic approach avoids the paradox by putting constraints on the usage of denial in complex sentences.

So, Priest explains that the acceptance and rejection rather depend on the evidence available: if we A and $\neg A$ are both well-supported, we have to accept both of them. Moreover, this explanation allows us to make sense of the incompatibility of acceptance and rejection: it comes from the understanding of the role evidence plays in the belief revision.

The minimal constraint on rationality is that we shouldn't accept everything: there are supposed to be rules that force us to accept some claims and reject the others. Usually, we tend to accept the claims (or at least we aim at this) that have good enough evidence for their truth. At first glance, it seems that we have an overall understanding of what good evidence is and check if the new information passes the test. If the statement A was justified by appeal to a dubious authority, nature or emotion, we don't judge the information as reliable. On the other hand, when the conclusion is based on a valid form of argument or induction with high probability, we are forced to accept it, even if it is in tension with the common sense. This picture allows for an easy explanation of why we accept or reject something, as there seem to be no difficulty in judging if the new information passed the test or not: the fallacy was committed or not, the argument was valid or not, the probability was high enough or not. It might happen that there is a piece of information A that passes the test, as well as the piece $\neg A$ – if evidence is good for both of them, we are to accept them both.

When considering inductive arguments we don't use the fixed threshold though, but rather compare the evidence for the competing hypotheses. When we compare evidence, we can find that although both of the hypotheses pass the "scientific plausibility" threshold, one is better supported, and so we accept it[28]. Furthermore, even though the evidence for the separate cases A and $\neg A$ may be quite good, their conjunction $A \wedge \neg A$ could be unsupported[29]. For instance, we could rationally believe that the contradictions involving observables are impossible and so suspend our belief about A and $\neg A$, because both of them are well-supported by evidence (Say, you see your friend on the street, but you know from a reliable source that she is not in the town (you just talked to her on skype, she showed you the view, etc.). You believe that she can't be in both places at the same time, so you suspend your belief before you come closer to see if you are mistaken or she had lied to you. If you had good evidence for the fact that people can be at the different places simultaneously, you could have just accepted both). In general, one has to accept A over B if $E_A > E_B$. Of course, it might happen that we have $E_A = E_B$; then we would accept both even if $B = \neg A$. We can't accept A and not accept it, because in this example it would mean that $E_A > E_B$ and $E_B > E_A$, which can't be the case.

This is the picture of rationality that can explain Priest's take on the "explosivity" of assertion and denial. For him, assertion is a linguistic expression of the cognitive act of acceptance. That is, one asserts B if and only if she accepts B[30]; dually for the denial and rejection. It means that denial and assertion are exclusive linguistic acts, as are acceptance and rejection. With

[28]Typically, much of the evidence for A is considered to be that the various alternatives to A ($\neg A$) are shown or believed to be wrong and hence rejected. In the paraconsistent framework this view has to be revised due to some counter-examples. Think of a Liar sentence: evidence for L doesn't depend on rejection of $\neg L$. In many cases the familiar view will apply though, if we know that the domains in question are consistent (i.e., we have enough evidence to think that $A \wedge \neg A$ is to be rejected).

[29]It has been suggested that logic of scientific inquiry and theory acceptance might be non-adjunctive (e.g., "high probability" logic) [Kyburg 1997; Teng 2011].

[30]Here assertion presupposes an environment free of lying, irony, sarcasm and other non-literal linguistic acts. In such a world truth and knowledge is the aim of asserting.

such a theory in hand Priest can answer the criticisms about the rationality of dialetheism, namely the question about possibility of disagreement with someone who accepts contradictions.

Although it is never sufficient to say $F(A)$ and $\neg T(A)$ if one wants to express disagreement or the fact that A is false only and can't be true, *denying* A would do the job, since it presupposes that the speaker can't correctly *assert* A on pain of being irrational[31]. If denial is understood as the *falsum*-negation, then the relation between the A-worlds and the $\perp$-world in the case of denial of A, i.e., assertion of $A \rightarrow \perp$, is intuitively understandable. If you accept and reject A, then you cease to be a rational agent, i.e., the one that has constraints on what is to be accepted and rejected; in such a case if one is to make everything out of this epistemological situation, she might as well find herself in the $\perp$-world, accepting anything.

One might object and say that such a conclusion would follow only if the

[31] I would like to make clear that, although denial is a solution to the *false only* problem, denying A is not equivalent to saying that it is *false only*, as we may intuitively understand it. Denial allows us to express genuine disagreement and formulate conditionals with dialetheias, which constitutes the "false only problem". Consider the example of disproving dialetheism from Chapter 1: there, from the fact the Liar sentence is not a dialetheia we wanted to conclude that dialetheism is *simply false*. It was impossible to express what we meant by this conditional if we understood the negation as negation of LP, since then negation of dialetheia is a dialetheia as well. Now we can reformulate this conditional in terms of denial to express what we really intended: "If the Liar is denied, then dialetheism is denied". Similarly, the problem of expressing genuine disagreement is resolved: dialetheist can just deny A if she disagrees with your claim that A; now you can be sure that she doesn't accept A as well. You won't automatically know what she thinks of $\neg A$, but dialetheist has no problems in providing you with this information (when she asserts A and denies $\neg A$, you know that she doesn't take A to be a dialetheia, since she can't assert $\neg A$ and deny $\neg A$ at the same time). What makes denial not equivalent to falsity *simpliciter* is the fact that one doesn't have to assert or deny A, for any A, since one can just withhold her opinion about the subject. In other words, "A is not *false only*" entails that A is *true* or *both true and false*, whereas if you do not deny A, you are not forced to assert A, since you might withhold your opinion on A. This might happen when you think that some propositions can't be accepted nor rejected (something unverifiable, for instance), or if you don't have enough evidence to decide. If one takes that there are no "undecidable" or "'unverifiable" propositions, i.e., that for every A, A is asserted or denied, then *false only* is equivalent to denial.

consistency of acceptance and rejection was the only constrain on rationality. But it is not true: on the picture presented, the evidence is a constraint. Our too tolerant agent will accept many things, but not all of them: there are claims that won't have the evidence that would make her accept them.

This objection is fair when one accepts the threshold view on evidence. It is true that if we have a threshold, say 50 args (measure of the amount of evidence: percentage of successful experience replication, probability of the hypothesis given the accepted theories, etc.), then everything that doesn't pass the test (everything that is < 50 args) will not be added to the agent's belief set and so the latter won't be trivial. I don't want to get into the discussion of what the correct view of the evidence is (threshold or comparative), but only point to the fact that the former requires the precise quantification of the evidence, which doesn't seem possible if the cognitive virtues are also taken into account (if they are not – one might accept absurd things into the belief set without constraints on *ad hoc* arguments). Rather, when one decides what to accept, she compares the hypotheses to the relevant rivals – the minimum of evidence can be sufficient to accept some statement, as long as it helps in understanding the given phenomena and preserves the already confirmed claims.

Does the objection go through if we take the comparative view on evidence? It doesn't. Take a random proposition A. When would an agent x accept it by a comparative view? When the evidence for it ($E(A)$) is better than for the relevant rivals ($E([B]_A)$, i.e., $E(A) > E([B]_A)$. What are the "relevant rivals"? They are definitely not the claims equivalent to A's contradictory, $\neg A$, as in the paraconsistent context it is safe to accept A and $\neg A$. They might be the claims that involve A's status of acceptance and rejection. We could be wondering if accepting A is better supported than A's rejection or than some claims that entail A's rejection. For instance, we might want to check if the evidence for the earth being round against the evidence for rejection of this claim: evidence for credibility of the space photography and officials presenting it, etc. The problem is that, if accepting and rejection are not exclusive, accepting the conditional $E(B) \Rightarrow Rej(A)$, where $Rej(A)$ means that A is to be rejected, doesn't force us to judge $E(B)$ as a "rival", as in the case of the contradictories, because both $Rej(A)$ (that implies $\neg Acc(A)$) and $Acc(A)$ might be true. In such case we are left with no evidence to compare with, so the notion of "rival" hypothesis looses its belief revision moving force.

The dialetheist is not someone who accepts all contradictions, but only some. We decide which ones are acceptable based on available evidence. There are some that can't be accepted given this evidence. For instance, the observational facts usually would be considered consistent, because there is no evidence to think otherwise (no Sylvan box has been found yet[32]). Semantic theory though gives enough evidence to think that the truth-predicate is inconsistent, but acceptable (based on induction on around two thousands years of thought about the Liar). By the same kind of reasoning, rejection and acceptance are contraries: the rational agent who accepts and rejects the same proposition is impossible, given the relation between acceptance and evidence. If the denial is analyzed as the *falsum*-negation, it makes the "explosion" of a set that contains the denied statement intuitively understandable, contrary to the criticism that says that it is "too strong".

[32]See [Priest 1997]. Priest wrote the fiction story to prove that one doesn't derive everything from a contradiction in certain contexts. In this story, when visiting Richard Routley's house he finds a box that contains a true contradiction in it. When you look inside the box you see that there is a statue of Buddha and it is not there at the same time.

To conclude: some thought that although introducing *falsum*-negation as the solution to the "false only" problem is an attractive option, for it is quite close to the familiar (from classical and intuitionist logic) notion, in the paraconsistent framework it loses its plausibility. In classical logic, where the notion of a maximally consistent and complete set (or the complete and consistent worlds [cf. Restall 1999]) serves as a basic framework, "either A is false, or triviality follows" seems to adequately express denial and negation of A (although one might criticise the adequacy of such an assumption about worlds). But in the dialetheic theory, they say, when one introduces the incomplete and inconsistent situations, and therefore the new conditional, the logical relation between the antecedent and the consequent used in the *falsum*-negation is too strong and thus unsuitable for the analysis of denial. This criticism is successful when a particular arrow used in the *falsum*-negation is presupposed (strict implication). Although better (for the dialetheic purposes) than the material conditional, such a conditional is not necessarily the most suitable one for the dialetheist. In this section I examined whether if using a ternary-relation conditional as the one governing the *falsum*-negation undermines the criticism. There are some plausible philosophical interpretation of the relevant arrows (namely, the understanding of the conditional as based on the relation of the relative relative possibility on information states) that make the *falsum*-negation look more attractive in the role of denial than was thought. This means that one might want to reconsider the sceptical attitude towards the *falsum*-negation as formalizing denial, especially given the problems that arise from "non-logical" approaches to denial. I will turn to evaluating them in the following chapter.

Chapter 3

Problems with the speech act solution to the *false only* problem

3..9 Introduction

As we saw, denial and assertion have to be exclusive for the dialetheist, who shares the common understanding of the relation between these linguistic acts and the cognitive acts of acceptance and rejection. Such a view in its turn forces one to distinguish denial of a statement and assertion of its negation on pain of triviality (otherwise, one would be forced to assert and deny the Liar at the same time, i.e., accept and reject it, which is impossible in the present view). Assertion and denial understood this way can solve the "just false" problem, i.e., can allow the dialetheist to express the genuine disagreement – just by denying a statement in question, instead of asserting its negation. In this view, denial is a notion that is not definable in terms of assertion. How is it defined then?

One would think that it could be adequately expressed by some kind of "just false" operator in the object language. In this case, the solution depends on how successful we are in defining such an operator. From chapter 1 we saw that this is quite a difficult task: as in the presence of such an operator the logic trivializes or requires further complications. Chapter 2 discussed the solution that came closer than others to achieving the desired result, but fell short of an adequacy constraint, from the point of view of some dialetheic theorists. This naturally led them to look for the solution elsewhere, in the extra-logical realm. Namely, Priest turned to speech act theory and analyzed denial and assertion as speech acts or force operators. I have argued that the *falsum*-negation is not necessarily a failure; in this chapter, I will provide the reasons to conclude that the speech act proposal has a more serious problem to deal with.

3..10 Assertion and denial as illocutory forces

By Priest's view (taken from [Parsons 1984]), assertion and denial are force operators, similar to question and command. The force operators are applied to the sentences' contents (propositions expressed by the sentence) and they show the aim of expressing the content and an attitude towards it. For instance, with the proposition $P =$ [The snow is white], i.e., the fact that the snow is white one can produce several sentences that express the same proposition, but in different ways. By adding the force of assertion (written as $\vdash$) to this proposition ($\vdash P$) or the force of questioning ($?P$), one gets sentences with different meanings, although the proposition expressed is the same in both cases[1].

In this view, meaning is not equivalent to the propositional content of the sentence, but rather is understood in a Gricean way [Grice 1957; 1968][2]. The difference in meaning in the case of $\vdash P$ and $?P$ stems from the difference in one of the conditions that have to be fulfilled for a successful speech act to happen. According to a simplified picture used by Priest, x asserts[3] p if she utters something for an audience A, with the intention that

(1) a certain response, r is produced in A,

(2) A recognizes that intention,

(3) the response to be produced in part by A's recognition of this intention [Priest 2006, p. 63][4].

In the case of $\vdash P$, the speaker x would intend that

(1) A accepts P or A beliefs that x accepts P,

(2) A understands that x is asserting P (not denying or commanding),

[1] One asserts A just in case when P is true ($\vdash A$ iff $1 \in v(A)$). A can be either just true or dialetheia. One can't make the similar condition for denial though. It is true that if you deny A, it is false, but not in the other direction, since you are not supposed to deny the sentences that are both true and false. Ripley [2011] claims that there is no way to state a condition for denial symmetric to assertion's one, unless we have a predicate Denial in our language (but that's bad, since then we can formulate a Liar. I think that the link between rejection and (comparative view of) evidence might help in formulation of such a condition. Deny A iff $E(A) < E(\neg A)$. For assertion: Assert A iff $E(\neg A) \leq E(A)$. Let's check if such condition works well. We know that we are supposed to assert the Liar, L, and its negation $\neg L$. We do so, because the evidence for both is the same, which suits our condition. In fact, it also explains what are the dialetheias in a minimalist way: we have good evidence for both of them (We might also add a constraint that the evidence for the conjunction should be high enough – in order to avoid an overproduction of dialetheias. By this view, if there is no evidence for either of conjuncts, it will be a dialetheia as well – there are no gaps for these cases, but gluts). This way the relation between hypothesis an its rivals is preserved as well: we do compare A and $\neg A$ in the case of denial. For instance, we will deny that the Earth is flat, since the evidence that it is not flat is better.

[2] Cf. [Davis 2011] for comparison of his early and later views. Priest refers to the simple picture of the early works.

[3] Here "asserts" means *produces a speech act with the content* p, not $\vdash p$

[4] "... to assert a contradiction is to behave in such a way as to try to get an audience to believe a contradiction, or at least to believe that the speaker believes it (by recognizing the speaker's intention to do just that" [ibid., p. 96].

(3) A accepts P or A beliefs that x accepts P, because A understands that it is what x intends[5].

3..11 How to know when the sentence is denied?

Consider the analysis of $?P$'s meaning. The conditions (2) and (3) are similar to $\vdash P$, and the (1) is: *the speaker x intends that A beliefs that x doesn't know if P is the case and that A asserts P or denies P.* Say, you ask me: "Is the snow white?" As your audience I understand that you don't know if the snow is white and that you suspect that I know it and will provide you with an answer. If I do know that it is white, I utter "Yes, the snow is white" with the force of assertion. If the speaker was pointing to a pile of snow recently marked by a dog, I'd utter "No, it is not white", i.e., deny that P ($\vdash^* P$).

One might suspect that the case of posing a question about a dialetheia is not so straightforward. Suppose I am asking: "Is the Liar sentence true?" ($?L$). A dialetheic member of the audience would say "Yes, it is true", i.e., $\vdash L$. But if she wants to be more informative and make her views on the topic clear, she'd say as well "The Liar is not true". The latter wouldn't mean $\vdash^* L$, but $\vdash \neg L$ – and neither are equivalent by the dialetheic view. Although the syntactical form of the negative answers to both questions is similar in English ("The snow is not white", "The Liar sentence is not true"), they are to be analyzed in a different way: "not" in the first case means denial of the statement (that the snow is white), whereas in the second case – assertion of the negation of this statement.

[5]Given the Gricean picture the link between assertion and acceptance is explained in the following way: the former is has as its aim the audience to recognize that the speaker accepts some proposition, i.e., the speaker has evidence for the truth of the proposition, and intends to produce the corresponding belief in the audience.

There is no easy answer to the question about how to know which speech act was produced (denial or assertion of negation). One has to derive the intentions of the speaker from the context, and here the "conversational implicatures" are supposed to point to what was meant – denial or negation [Priest 2008, Ch. 6]. For instance, when a person looks at a picture, frowns and says with a sarcastic intonation "What a beauty...", we would understand that she meant to deny the proposition that the picture is beautiful.

Although this seems to be convincing, there are some examples that are not so easy to deal with, the ones where there is no context from which one could derive the intentions of the speaker. Consider an example closer to our topic:

We are in a classroom with students waiting for Priest to come and give a lecture. One of the students enters the room late and asks: "Is the professor is already here [in the room]?" Another student, Olga, who sees Priest in the doorway, answers "No, he is not in the room yet". Priest hears the answer and hurries to disagree with Olga, as he is persuaded that when you are in the doorway, you are in the room and you are not in the room at the same time [Priest 2006, Ch. 12]. What could he say about Olga's utterance O in order to express his (corrective) disagreement[6]?

1. "O is false". This is supposed to convey his intentions as he doesn't agree with Olga that "he is not here yet", because he is (and not). But Olga would most definitely think that Priest means that he is already in the room (full stop) if she is not already aware of Priest's theory – and nothing else from the context would make her think differently. If Priest didn't inform his students about his beliefs about the doorways before, he has to be more

[6]For he doesn't want to discard O, but rather add information to it

specific, as the maxim of informativity requires;

2. "O is true" results in the same situation;

3. "O is both true and false" seems to be what Priest needs, as it expresses fully his thoughts on the topic.

So, suppose he utters "O is both true and false". Did he disagree with Olga by doing so? If we decide that in order to express disagreement one has to say that somebody's utterance is false, and to express agreement one has to say that it is true, then he did disagree with Olga (because he said that O was false) and he didn't (because he said that it was true as well). In this case if Priest decided to express his disagreement with Olga by saying "I disagree with Olga", which is equivalent to 1., it wouldn't be informative enough, and he would have to say "I agree and disagree with Olga". Given that the students are probably aware of the peculiarities of their teacher, probably the communication wouldn't fail. In fact, it may be a good consequence, as one is talking about the corrective disagreement – it is not supposed to be exclusive – the disagreed with statement can be asserted with the correction statement ("It is a beer and it is the best beer in the world!", "John is willy or nilly. He is willy and nilly", etc.).

Now imagine a situation where Priest is in the hall, not in the doorway yet, and another student, Mary, answers the initial question by saying (M): "He is in the room and not in the room" (since she already knows Priest's theory and mistakenly thinks he is in the doorway). Priest sees that Mary is mistaken as he is one step before the doorway, definitely out of the room. So, what should he say if he wants to correct the student?

Probably, "M is false". But then he would just confirm Mary's answer (as in the case were he would say "M is true"). What he intends to say is that it is *not both true and false*, and, moreover, he has to do that in order to be as informative as possible. By formal (paraconsistent) logic alone we know that this sentence is equivalent to saying that "M is *both true and false*" (Both $=$ $\neg$Both).

But he didn't mean that. Here, by "not" he signalled that he *denies* the sentence in question (M), i.e., genuinely disagrees with it. It might seem similar to the case where a person uses an implication – most probably she doesn't mean to use the material implication, required by formal (classical) logic. This makes us come back to Grice's treatment of logical connectives and "conversational implicatures" involved. For instance, the conjunction would be interpreted in a different from the formal logic way. We would have $A \wedge B$ is true iff A is true, B is true, and $t_A \leq t_B$, where t_X means the time when X happened. (This explains the difference in meaning of the classical examples: "She married and got pregnant" and "She got pregnant and married" – the truth conditions will be different. The commutative cases will involve non-actions, and we can postulate for them that $t_1 = t_2$, where 1 is the leftmost conjunct and 2 is on its right. "The grass is green and trimmed" would be true as well as "The grass is trimmed and green". The conjunction of the different types "The grass is green and was trimmed" would be commutative as well.) As long as we have the appropriate context, we can rephrase or explain what the speaker really meant. So, in some cases the usual connectives would mean something different, although written or expressed in the same way, as instead of the usual commutative "and" $\wedge$ in the ordinary language we will have a new $\wedge^*$ with the time constraint.

But in the above examples, there seem to be no "context" to help us to know if "false" or "not" that are used meant genuine disagreement or the corrective one, except when we already know that our interlocutor is a dialetheist who would be willing to assert Liar and its negation (correctively disagree that the Liar is false), but deny neither of them (not genuinely disagree that the Liar is false). This in turn means that we already have to know what the speaker meant, i.e., what she intended, but that's exactly what the context was for! In order to ameliorate the situation we could introduce the convention of saying explicitly when we deny a statement and when we assert it (In English it is common to answer an inversion question "Is it?.." by saying"Yes, it is..." or "No, it isn't..." and it is presupposed that these answers are incompatible, although one could say "Yes, it is..., but also...", where "..." stands for the contradictories). The answer to $?P$ then would be "I deny P (and assert $\neg P$)" and the answer to $?L$ would be "I assert L and $\neg L$ (but not deny L)".

Now we will see that this convention doesn't work well for this view.

3..12 Embedded denial

The nice feature of the speech-act formalism is that it allows us to avoid the extended Liars. I will take this conclusion for granted and discuss the move Priest has to make in order to achieve this (that the force operators have no interaction with the content), for it reveals an important problem the present approach has. This problem will make us better appreciate the *falsum*-negation solution.

The sentence "I deny that this very sentence is true" doesn't create a problem. Take the usual Liar sentence, L, equivalent to $\neg T(L)$, then the new Liar will be $\vdash^* T(L)$. Even though we get $\vdash^* \neg T(L)$ by substitution, it is not a problem for a gap theorist, because it is exactly what is supposed to be: she says that it doesn't have a truth value. It doesn't create a problem for a dialetheist either, as it is just a denial of the truth-teller. It might be wrong to do that, but not paradoxical. These examples show that "attempts to formulate distinctive Liar paradoxes in terms of denial fail, since, being a force-operator, has no interaction with the content of what is uttered" [Priest 2008, p. 104].

What about a more complicated sentence: "I deny that this very sentence is denied" [Parsons 1984]? Priest's answer to this Liar is to say that it doesn't make sense to use denial in embedded contexts, because it is a force operator and can't be applied to a force operator, only to the content (proposition). The statement, "Is the sentence A denied?", doesn't really mean what it seems to mean ($? \vdash^* A$): one can't in the same breath wonder about and deny a certain proposition. In any case, that's not what the above statement means (rather, one intends to check the denial-assertion status of the proposition A). But previously we agreed to mean by "deny" $\vdash^*$, because otherwise we are not able to communicate our intentions, as the "context" is not informative enough for some cases. The (impossibility of) embedded use of denial makes us reconsider this convention.

Some less tricky sentences are not easy to analyze within this framework either. Consider the sentence "If you deny A, then you deny B". If "deny" is to be understood as a force operator in this case, then we have two propositions here [A] and [B] and the forces of denial applied to them: If $\vdash^* A$, then $\vdash^* B$. The question is how this sentence is to be analyzed? We are certainly not denying the *if...then* clause. More plausibly we are asserting the proposition expressed by the sentence (condition of it being a sentence at all). Let call this proposition [If...then]. Then we have $\vdash$[If...then]. But the members of

[If...then] count among themselves $\vdash^* A$ and $\vdash^* B$, which can't be the case, because one can't apply two forces to the same proposition (moreover, it is not what was being said: "I assert and deny that if A, then B"). But if it applies only to proposition expressed by the parts of the sentence, then we get $\vdash$[A],[B], which is not what the given sentence means. What we want here is that $\vdash$ is applied to the content of denying A and B ($\vdash [\vdash^* A \Rightarrow \vdash^* B]$, which is not well-formed). But, this can't happen, according to Priest, since denial is to be analyzed as a force operator.

We can still preserve our earlier convention and posit that when embedded "deny" is to be analyzed differently, as a part of the content, for instance. This allows to solve the issue of the communication failure, but creates another problem, which in the literature related to philosophy of language is known as the "Frege-Geach" problem.

3..13 The Frege-Geach problem

The problem with the speech act account of denial I am going to discuss is already familiar from the critique of "expressivism" or "noncognitivism" in moral philosophy[7]. It might seem surprising that the problem is parallel to the one found in the sphere of semantic paradoxes, but the explanation of this fact is quite simple: both theories try to divide the words of the natural language into two categories: for those terms that have the usual truth-functional, compositional meaning and the others whose meaning is expressing an attitude towards the content.

The clear-cut examples of the latter, from the "expressivist" point of view,

[7] The most well-known version of it presented by Ayer [1936] and Carnap [1935] against metaphysics and ethics as its subspecies. Hare [1952] developed the view. See [Schroeder 2008] for a good summary.

are such "alleged" adjectives as "good", "just", "wrong", etc. They are to be distinguished from the observational terms, e.g., "green", "11,000 feet high". Although they are both adjectives (or adverbs) and seem to play the same role in the sentences they form, their meanings are to be analyzed in a different way.

Compare the phrases "The grass is green", "This action is good"; "This image is not green", "Murdering is not good". "Good" and "green" seem to ascribe a property to an object the subject of the sentence stands for, i.e., the truth-conditions of both sentences seem to be of the similar form. "The grass is green" is true iff the grass is green; "this is good" iff this action is good. These simple considerations lead us to thinking that "good" corresponds to some fact about the world in the same way as "green" does.

This is precisely the view that the expressivists are criticizing. According to them, the language misleads us, and despite the appearances the "good"-type terms (moral and metaphysical notions) have different semantics from the "green"-type ones. The former doesn't have the same truth-conditions as the latter. It is incorrect that "Murdering is wrong" is true iff murdering is wrong, because the words like "good", "just", etc., are not predicates and don't classify objects, in fact, they are the expressions of command or desire. When someone says that murdering is wrong, she doesn't mean to say that the act of murdering has the property of "wrongness", rather she is expressing the command "Do not murder!" or her desire "I wish there were no murdering!".

The "Frege-Geach" criticism of the above account amounts to showing that there are cases of the usage of the moral terms that can't be incorporated in this system, but that are highly intuitive. For instance, it seems that although it might be difficult to prove the truth of the premises of a moral argument, we are still able to correctly reason about their consequences. If we established that *if murdering is wrong, then stealing is wrong* and that in fact murdering is wrong, we would be forced to conclude that *stealing is wrong.* In order for this to be a valid application of Modus Ponens, and it seems to be the one, the meanings of the antecedent in the first premiss and the second premiss have to be the same; otherwise, we would commit the fallacy of equivocation. So, we seem to be able to have the embedded use of the moral terms; but if we do, then the embedded terms have to have the same meaning. The problem for the expressivist is to show how the force-operator can indeed be embedded within the sentence.

The same applies to dialetheist if denial and assertion are analyzed as force operators. If one is committed to say that analyzing denial as a force operator is correct, that is, it explains correctly the semantics of the latter (despite the appearances), then saying that it can't be embedded in the sentence, as Priest does in order to avoid the extended Liars, is ignoring the important usage of the terms and suggesting that we can't reason about the consequences of what we accept and reject. The sentence "I deny the fact that this sentence is denied" makes sense in English and the dialetheist who subscribes to a speech act view of denial is obligated to explain how it is a violation of the semantic rules governing denial. It makes sense as well to say (remembering that denial is a solution to the "false only" problem): "If you deny A, you are to deny B". If one is not allowed to reason in this way, then it is doubtful that it is a good analysis of the notion (because it ignores the highly intuitive cases) and but also that it is a useful notion (if one can't use but for describing atomic sentences).

3..14 Possible answer to the criticism

Now I would like to see if the dialetheist could use some solutions to the Frege-Geach problem already proposed in the moral philosophy literature. The central idea of these solutions is to give compositional semantics for speech act terms [Schroeder 2008, p. 708]. Such an approach is not suitable for Priest, since for him it was exactly the appeal to the non-compositionality of force operators of denial and assertion that provided a Liar-free solution to the *false only* problem. Hence, I will discuss another way of answering the Frege-Geach criticism: redefining the notion of validity in terms of higher-order attitudes.

Since the main point of the Frege-Geach criticism was that expressivism contradicts the fact that the moral sentences can be embedded in the complex sentences (which is proved by validity of moral arguments), the aim of the late generations of expressivist theorists was to explain how moral terms understood as force operators could be in fact embedded (for instance, how one could use *Modus Ponens* with moral sentences as premisses) [Hare 1970][8]. In other words, they tried to provide the compositional semantics for non-descriptive terms that would do the same thing as the descriptive terms.

> This is the new shape of the Frege-Geach Problem, and it is the one that noncognitivists have been trying to address since Hare. The problem is to construct a compositional semantics for natural languages which makes complex moral sentences and complex descriptive sentences turn out to have the same kinds of semantic properties - and the right kind of semantic properties - even though moral and descriptive terms really have two quite different kinds

[8]More recently Gibbard [1990; 2003], Horgan and Timmons [2006].

of meaning.

[Schroeder 2008, p. 706]

Priest can't use Hare's answer and say that the non-truth-functional terms behave in a similar way to the truth-functional terms, that denial put into conditional has the same meaning as when it stands independently (what is required to explain the validity of *Modus Ponens*), because it makes the same contribution toward the truth-conditions of the whole sentence. This is so, because for Priest there is another constraint – the Liars shouldn't reappear. Postulating the difference in meaning for embedded denial (i.e., prohibit self-reference) was a way to avoid extended Liars.

Alternatively he could challenge a presupposition of Geach's criticism – that in order for *MP* to be valid the terms in the premisses and conclusion have to mean the same thing. Since arguing that these terms do in fact have the same meaning would lead to paradoxes in Priest's view[9] he could argue that the validity of inferences involving force operator terms (assertion and denial) is explained in a different way.

Such a solution to the "new shape of the problem" in the moral philosophy is known as an appeal to the "higher order attitudes" (Blackburn [1984]). Priest could try using it, since the expressivist's analysis of moral terms is quite similar to his approach to denial and assertion. Recall that in the expressivist view, the sentence of the type "This is wrong" expresses the mental state of *disapproval* of the referent of "this". And the expression of this state was the command: "Don't do that!" Similarly, for Priest, denial expresses rejection

[9] Another way of avoiding the paradox is to make constraints on logics, so that the Liar is not derived. In this case we come back to the "object language" solutions to the paradox, e.g., the *falsum*-negation.

attitude towards the contents of the sentence. The application of denial as a force operator to the content means: "Reject this sentence!" (recall what constitutes the meaning of the speech act of denial: to produce in the speaker the belief that the content is rejected).

The attempts to solve the Frege-Geach problem from the expressivist point of view aimed to explain why *MP* is valid in the terms of attitudes. Simon Blackburn [1984], for instance, talks about the hierarchy of attitudes. In his view, conditionals express higher-order attitudes toward attitudes expressed by their parts. For instance, "If murder is wrong, then stealing is wrong" expresses disapproval of the state of both disapproving of murder and not disapproving of murder, i.e., it expresses a higher-order attitude toward the mental states expressed by the parts of the sentence [Schroeder 2008, pp. 708-710]. This is supposed to explain why we can reason in moral terms: we are obliged to accept the conclusion on pain of moral incoherency (disapproving of our own actions). By expressivists' view, every conditional involving moral terms can be reformulated in terms of cognitive acts: when one accepts the premises but not the consequences she is in a situation she is disapproving of. So, the conditionals using the moral terms are explained in moral terms as well. This way although *MP* is shown to be valid, but it is valid not in virtue of syntactic properties of sentences, but because of mental states they express.

It seems that one could as well use this kind of approach to explain the validity of *MP* involving the speech acts of assertion and denial, since they are analyzed in terms of cognitive acts too. Moreover, we already agree that one can't reject and accept something at the same time (the incoherency involved can be explained in terms of belief revision [see section 2.8 of the present work]). For example, consider a conditional: (1) *If A is denied, then B is* *denied*. In terms of cognitive acts, this sentence would be analyzed as follows: (2) *That A is rejected and B is not rejected is rejected.* (2) is supposed to explain why inference from (1) and (3) *A is denied* would lead to (4) *B is denied*: we can't accept (rejection of A and B) something we reject.

Unfortunately, this approach to solving the Frege-Geach problem[10] reintroduces the Liar paradox, since now the analysis of denial becomes equivalent to the "object language" approach (you introduce in the language the predicate $Rej(x)$ to mean *false only*).

(2) has two appearances of "rejected". If we say that they mean the same thing, it would mean for our semantics of denial that we can construct a self-referential sentence: "Reject that this sentence is rejected". If you reject this sentence, then it is true that this sentence is rejected, but then you have to assert it, since you have to assert true sentences. But if you assert it and its true, then it is rejected as well[11].

To avoid self-referential problems, we could postulate the hierarchy of attitudes and say that in (2) "rejected" doesn't express the same attitude on every appearance. The last appearance of "rejected" will express the higher-order attitude towards the rejections within the sentence. This might be an option for dialetheist, since it helps to avoid the Liar, but it is implausible as an analysis of how language works (this was the point of the Frege-Geach problem in the first place). Consider two conjunctions: (5) "Deny A and B" and (6) "Deny A and deny B". (5) is analyzed in terms of rejection in the following

[10] There are reasons to think that it doesn't solve this problem. Namely, that such interpretation of validity might over-generate validity [Schroeder 2008, p. 709].

[11] Koons [1992, pp. 88-89] derives a Liar using only 4 simple rules that might govern the acceptance and rejection if analyzed as predicates like Truth in the embedded contexts: (1) Rej(s) $\rightarrow$ Proved(Rej(s)); (2) (x) (Proved(x) $\rightarrow \neg$ Rej(x); (3) From $\vdash (\phi \rightarrow \neg\phi)$ derive $\vdash^{*}\phi$; (4) From $\vdash^{*}\phi$ derive Rej(x); compare to [Ripley 2011] who talks about inevitability of the *Denier* paradox in "object language" frameworks.

way: (5a) *Reject A and B*, whereas (6), analogous to (2), – as (6a) *Reject that A or B is rejected*, since denial is embedded in (6). The implausibility of such view becomes evident when we consider assertion (7) *Assert that A and B is asserted*. By the present view, it is not equivalent to (8) *Assert A and B*, because "assert" in (7) expresses higher-order attitude towards the contents. We could postulate that assertions of different levels are equivalent as they seem to be, but still keep separate the rejections of different levels. But such a solution would lead back to the Liar paradox: consider a case where you reject$_{n+1}$ the sentence "This sentence is rejected". You assert it on level n, but don't reject$_n$ it, which helps avoiding the paradox. But if the assertions are the same on different levels, then assert$_n$ = assert$_{n+1}$ and we arrive at another "incoherency". One could approach carefully the construction of such an hierarchy in order to avoid this problem, but it doesn't seem to suit for dialetheist, since saying that the Liar sentence is *both true and false* (which caused the *false only* problem) was motivated by attempt to avoid "artificiality" of Tarski's hierarchy of Truth-predicates.[12].

[12] See [Bremer 2005, Priest 2006]. There is a better way to make denial compositional that don't involve such complications; I don't consider it here, since it has been established that it doesn't avoid Liar either. [Restall 2008; 2013] uses multi-conclusion sequent calculi tools in order to analyze denial. $S_1 \vdash S_2$ means that denying S_2 and asserting S_1 is irrational. Given what one takes to be the logical laws governing assertion and denial one can build a logic of denial and assertion using such a structure. It is general enough to avoid the problems involving arrows in non-classical settings, but also express embedded denial. We take $\vdash$ to be reflexive, transitive logical consequence relation ("coherence"). The empty set of premises gives us a guide to what accept on the basis of the logic alone. As we remember denial and assertion are supposed to be exclusive. The former means that we will have the rule $A \vdash A$: it is incoherent to deny and assert A; denial is not equivalent to the assertion of negation: $\vdash \neg$ (means that not-A is undeniable (a theorem)), then $A \vdash$ (means unassertable (falsity)) can't be a rule of logic. The dual rule is defines though $A \vdash$, then $\vdash \neg A$ (If A is unassertable, then not-A is undeniable). X $\vdash$ Y means that it is incoherent to deny each member of Y and assert each member of X. Let Γ and Δ will be the sets of sentences and A an individual sentence. Then, by Restall denial as understood by a dialetheist is to be understood as governed by the following rules

(Identity) $\Gamma, A \vdash A, \Delta$

(Cut)

3..15 Conclusion

In this chapter I discussed the following problem for the speech-act approach: if one decides to analyze denial as a force operator, the question arises how to know when this operator is applied and not others (questions, commands, assertions, etc.), as "not" can be ambiguous between its "corrective disagreement" and "genuine disagreement" meanings. In the abstract (think: written by an anonymous writer) case, the context is not helpful in determining which meaning was used, so the "conversational implicature" view doesn't work. We could have postulated a convention to analyze the word "deny" as expressing

---

$$\frac{\Gamma \vdash A, \Delta \qquad \Gamma, A \vdash \Delta}{\Gamma \vdash \Delta}$$

($\neg$-Introduction)

$$\frac{\Gamma \vdash A, \Delta}{\Gamma, \neg A \vdash \Delta}$$

(Truth-schema)
(LT)

$$\frac{\Gamma, A \vdash \Delta}{\Gamma, T\langle A \rangle \vdash \Delta}$$

and (RT)

$$\frac{\Gamma \vdash \Delta}{\Gamma \vdash T\langle A \rangle \Delta}$$

$$\frac{\Gamma, T\langle A \rangle \vdash T\langle B \rangle, \Delta \qquad \Gamma, T\langle B \rangle \vdash T\langle A \rangle, \Delta}{\Gamma \vdash \langle A \rangle = \langle B \rangle \Delta}$$

$\langle A \rangle$ can be either truth-value or a proposition expressed by A.
(=L)

$$\frac{\Gamma, \phi(a) \vdash \Delta}{\Gamma, a = b, \phi(b) \vdash \Delta}$$

[Restall 2013, pp. 94-98] derives a Liar paradox using these principles, and gives the homework for the future: "a friend ... of the T -scheme for truth must explain which of (Id), (Cut), (=L) and (Ext & Int) are to be rejected..." See [Ripley 2013; Cobreros et al. forthcoming] for a proposition to dispose of Cut, for instance.

denial and not assertion of negation. But this is incompatible with the view that denial can't be embedded in the sentences – as there are clearly some cases when it can, but then "deny" can't be analyzed as a force operator. If we decide to give up on the previous convention, we end up in a situation, where we have troubles understanding what was said, as the context is not telling, but, furthermore, in a situation, where we are forced to say that the reasoning about denial is never valid (the Frege-Geach problem). Compositional semantics for these terms are unavailable, since then the Liar could be reformulated. What one could do is make restrictions on such semantics that would prevent the formulation of the Liars.

But the latter reply was already available on the "object language" level, so it seems to be redundant to appeal to the speech acts, if they lead to the even more complicated consequences. In Chapter 2 I explained that *falsum*-negation doesn't pose a philosophical problem that was attributed to it, and it is generally agreed that it could avoid the Liars, so it seems to be preferable to the speech-acts approach.

Of course, the latter claim has to undergo further examination in order for this solution to be accepted. In particular, the various extended Liars have to be tested[13]. Typically, the inferences of the Liar for the *falsum*-negation are invalidated due to the failure of *Contraction*. One has to show how this failure interacts with embedded denials and valid *Modus Ponens*[14] In this respect

[13][Ripley 2011] suggests that there still might be troubles in expressing disagreement when denial is "gappy" – this is a point to check further. By the theory of *falsum-negation* you want to deny that $A \vee (A \rightarrow \perp)$, i.e., that every A is either asserted or denied. Say, A is equivalent to a "Denier", a sentence that says of itself that it is denied $\delta \leftrightarrow (\delta \rightarrow \perp)$. You would want to disagree that A is asserted. But you can't express it using $(\delta \rightarrow \perp)$, since then you would assert δ. If this really true depends on particularities of our semantics for conditional.

[14][Beall 2009] uses the division on normal and abnormal words – *Modus Ponens* is valid on normal points and *Contraction* fails on abnormal points. [Priest 2006; Restall 1996]

Restall's approach to denial is promising (see the footnote 11), since it doesn't involve appeal to a conditional. Still, it faces the similar problem – Liars can be derived.

In this thesis I didn't talk about other "extra-logical" solutions to the *false only* problem. In particular, I didn't talk about JC Beall's distinction between the rules of rationality and logic. He proposes to solve the logical problems (invalidity of *MP*, the *false only* problem) by accepting that logic tells us only as much – we have to choose from the alternatives provided by logic on some different rational grounds. There is another approach that Ripley [2011] calls paracoherentism: to accept that we deny and assert certain cases simultaneously. Similarly, Priest talks about irresolvable rational dilemmas [Priest 2008, Ch. 7]. We are supposed to manage these logical problems by common sense, rationality, etc. These approaches have different philosophical presuppositions (in particular, about the relation of logic and rationality), so in order to criticize them one would have to do a detailed philosophical analysis of their position. Rather, I am showing that there is a possibility of solving the problem within logic[15] and that the philosophical problem associated with it can be solved. I leave for further research to develop the *falsum*-negation solution to the *false only* in its fullest extent.

provides the restrictions on relations. These are the frameworks to compare in order to come up with a suitable interpretation.

[15] In particular, for [Beall 2009; 2013], the failure of logical laws in paraconsistent logics shows that one has to appeal to extra-logical rules

Bibliography

Armour-Garb, B., Beall, J.C. [2005] *Deflationary Truth*, Chicago, Open Court.

Armour-Garb, B., Beall, J.C. [2005] *Deflationism and Paradox*, Oxford University Press.

Armour-Garb, B., Beall, J.C., Priest, G. [2004] *The Law of Non-Contradiction*, Oxford University Press.

Ayer, A. J. [1952 (1936)] *Language, Truth, and Logic*, New-York, Dover Publications.

Beall, J.C. [2013] "A Simple Approach Towards Recapturing Consistent Theories in Paraconsistent Settings", *The Review of Symbolic Logic*, online publication (August 2013), Springer.

Beall, J.C. [2013a] "Shrieking Against Gluts: the Solution to the 'Just True' Problem", *Analysis*, 73 (3).

Beall, J.C. [2013b] "Free of Detachment: Logic, Rationality, and Gluts", *Nous*, 47 (3).

Beall, J.C., Ross, B., Dunn, M., Hazen, A., Mares, E., Meyer, R., Priest, G., Restall, G., Ripley, D., Slaney, J., Sylvan, R. [2012] "On the Ternary Relation and Conditionality", *Journal of Philosophical Logic*, 41(3) (June 2012), pp. 595-612.

Beall, J.C., Ripley, D. [2011] "Nonclassical Theories of Truth", in Glanzberg, M. (Eds.), *Oxford Handbook of Truth*.

Beall, J.C. [2009] *Spandrels of Truth*, Oxford University Press.

Beall, J.C. [2007] "Prolegomenon to future revenge", in Beall, J.C. (Ed.) *Revenge of the liar: new essays on the paradox*, Oxford University Press, 2007.

Beall, J.C. [2006] "Truth and Paradox: a Philosophical Sketch", in Gabbay, D., Thagard, P., Woods, J., *Handbook of the Philosophy of Science*, pp. 186-272.

Beall, J.C., Brady, R., Hazen, A., Priest, G., Restall, G. [2006] "Relevant Restricted Quantification", *Journal of Philosophical Logic*, 35 (6), pp. 587-598.

Bremer, M. [2005] *An Introduction to Paraconsistent Logics*, Frankfurt, Peter Lang Editions.

Bromand, J. [2002] "Why Paraconsistent Logic Can Only Tell Half the Truth", *Mind*, 111 (444), pp.741-749.

Carnap, R. [1935] *Philosophy and Logical Syntax*, Bristol UK, Thoemmes.

Cobreros, P., Egr, P, Ripley, D., van Rooij, R. [forthcoming] "Reaching Transparent Truth", *Mind*.

Copeland, B.J. [1979] "On When Semantics is Not Semantics: Some Reasons for Disliking the Routley-Meyer Semantics for Relevant Logic", *Journal of Philosophical Logic*, 8, pp. 399-413.

da Costa, N. C. A., Alves, E. H. [1977] "A Semantical Analysis of the Calculi C_n", *Notre Dame Journal of Formal Logic*, 18 (4), pp. 621-630.

Davis, W. [2011] "Implicature", *Standford Encyclopedia of Philosophy*, `http://plato.stanford.edu/`.

Everett, A. [1993] "A Note on Priest's 'Hypercontradictions'", *Logique et Analyse*, volume 141-142, pp. 39-43.

Fitch, F.B. [1946] "Self-reference in Philosophy", *Mind* 55, pp. 64-73.

Field, H. [2008] *Saving Truth From Paradox*, Oxford University Press.

Gabbay, D., Guenthner, F. [2002] *Handbook of Philosophical Logic*, vol. 6, Dordrecht, Springer.

Geach, P. [1965] "Assertion", *Philosophical Review*, 74, pp. 449-465.

Grandy, R., Warner, R. [2011] "Paul Grice", *Standford Encyclopedia of Philosophy*.

Grice, P. [1968] "Utterer's Meaning, Sentence Meaning and Word Meaning", *Foundations of Language*, 4, pp. 225-242.

Grice, P. [1957] "Meaning", *Philosophical Review*, 66 (3), pp. 377-388.

Hare, R. [1952] *The Language of Morals*, Oxdford University Press.

Horn, L. [1989] *The Natural History of Negation*, University of Chicago Press.

Koons, P. [1992] *Paradoxes of Belief and Strategic Rationality*, Cambridge University Press.

Kyburg, H. (Jr) [1997] "The Rule of Adjunction and Reasonable Inference", *The journal of Philosophy*, 3, Volume 94, pp. 109-125.

Mares, E. [2012] "Relevance Logic", *Standford Encyclopedia of Philosophy.*

Mares, E. [2004] *Relevant Logic: A Philosophical Interpretation*, Cambridge University Press.

Parsons, T. [1984] "Assertion, Denial, and the Liar Paradox", *Journal of Philosophical Logic*, 13(2), pp. 137-152.

Priest, G. [2013] "Vague Inclosures", in Tanaka, K., Berto, F., Mares, E., Paoli, F. (Eds.) *Paraconsistency: Logic and Applications*, Series: Logic, Epistemology, and the Unity of Science, Vol. 26, Springer.

Priest, G. [2008] *Doubt Truth to Be a Liar*, Oxford University Press.

Priest, G. [2006] *In Contradiction*, Oxford University Press.

Priest, G. [2002] *An Introduction to Non-classical Logic: From If to Is*, Cambridge University Press.

Priest, G. [1997] "Sylvan's Box: A Short Story and Ten Morals", *Notre Dame J. Formal Logic*, 38 (4), pp. 573-582.

Priest, G. [1990] "Boolean Negation and All That", *Journal of Philosophical Logic*, 19(2), pp. 201-215.

Restall, G. [2013] "Assertion, Denial and Non-Classical Theories", in Tanaka, K. et al. (Eds.) *Paraconsistency: Logic and Applications*, Dordrecht, Springer, pp. 81-100.

Restall, G. [2008] "Assertion, Denial, Accepting, Rejecting, Symmetry, Paradox, and All That", forthcoming, Melbourne, `http://consequently.org/writing/assertiondenialparadox/`.

Restall, G. [2005] "Multiple Conclusions" in Hajek, P., Valdes-Villanueva, L., Werterstahl, D., *Logic, Methodology and Philosophy of Science: Proceedings of the twelfth International Congress*, Kings College Publications, pp. 189-205.

Restall, G. [1999] "Negation in Relevant Logics: How I Stopped Worrying and Learned to Love the Routley Star ", in Gabbay, Dov M., Wansing, H.(Eds.), *What is Negation?*, Dordrecht, Kluwer Academic Publishers, pp. 53-76.

Ripley, D. [2013] "Paradoxes and Failures of Cut", *Australasian Journal of Philosophy*, 91(1), pp. 139-164.

Ripley, D. [2011] "Embedded Denial", `https://sites.google.com/site/davewripley/papers`.

Ripley, D. [2011a] "Negation, Denial and Rejection", *Philosophy Compass*, 6(9), pp. 622-629.

Searle, J. [1962] Meaning and Speech Acts", *Philosophical Review*, 71, pp. 423-432.

Schroeder, M. [2008] "What is the Frege-Geach problem?", *Philosophy Compass*, 3(4), pp. 703-720.

Smiley, T., Priest, G. [1993] "Can Contradictions Be True?", *Aristotelian Society*, Supplementary Volume 67, pp.17-54.

Tappenden, J. [1999] "Negation, Denial and Language Change in Philosophical Logic", in Gabbay, Dov M., Wansing, H.(Eds.) *What is Negation?*, Dordrecht, Kluwer Academic Publishers, pp. 261-298.

Tarski, A. [1983 (1959)] "The Concept of Truth in Formalized Languages", in Corcoran, J. (Ed.) *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*, Oxford University Press.

Teng, C. M. [2012] "When Adjunction Fails", *Synthese*, 186(2), pp. 501-510.

Woods, J. [2003] *Paradox and Paraconsistency: Conflict Resolution in the Abstract Sciences*, Cambridge University Press.