Paraconsistent Logics and Identity
- a Pragmatic Approach

MSc Thesis *(Afstudeerscriptie)*

written by

Thomas Cattermole
(born August 27th, 1987 in Colchester, UK)

under the supervision of **Prof Robert van Rooij**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:	**Members of the Thesis Committee:**
August 31, 2016	Dr Maria Aloni
	Prof Franz Berto
	Prof Benedikt Löwe (chair)
	Prof Robert van Rooij

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

This thesis investigates what role paraconsistency can play in dealing with traditionally intractable problems concerning identity and change. More specifically, we consider three logics that all commit to a Leibnizian account of identity, but that provide distinct solutions to a version of the sorites paradox related to the Ship of Theseus. It is shown how the first of these logics solves the problem by taking an approach that embraces inconsistency, rendering invalid some familiar principles, while the second takes a more consistent approach, satisfying these familiar principles, but is not able to capture some core elements of paraconsistency. The final logic uses positive features of both approaches to provide a radically different account of the nature of change. This logic is particularly interesting in that it is a non-monotonic logic inspired by pragmatic considerations regarding vagueness. The viability of each approach is assessed in part according to their treatment of the paradox, and in part according to five rules of inference that we argue any account of identity should allow.

Acknowledgements

Firstly, I would like to thank Robert for being an excellent supervisor. I was able to learn a lot quickly through his guidance and I always enjoyed our meetings. I am particularly grateful for the ever prompt and perceptive responses via email while I was writing this thesis away from Amsterdam. I am also grateful to my defense committee for their feedback, including the suggestion that I include the code which is now present in the appendix. My thanks also go to all those at the ILLC responsible for making the MSc in Logic the challenging and stimulating programme that it is. A special mention should go to Tanja who always helped me whenever I needed it. Thanks should also go to my fellow students, in particular to Marcus for our countless debates and discussions. I would also like to thank Madalina for her love and support throughout my studies, and especially during the writing of the thesis. The sharp proofreading given of final drafts made the writing better than it would have been otherwise. Lastly, I would like to thank my family and in particular my parents for the financial as well as emotional support that made this all possible.

Contents

Chapter 1

Introduction

Doubts concerning the principle of bivalence - that every statement has exactly one truth value, either true or false - are as old as logic itself.[1] It was not until the early Twentieth Century and the independent work of Łukasiewicz [21] and Post [24] that such doubts could find a place within formal systems: the arrival of Many-Valued logics. Since then, numerous such logics have been proposed, and with them as many purported applications, some purely philosophical, others less so. For our present purposes, the most important of these contributions have been those capable of making some sense of the issue of borderline cases, i.e. vagueness. Might it be that some propositions are vaguely true in the sense that they are neither true nor false (i.e. truth-gaps)? Or perhaps, there are some propositions that are both true *and* false (truth-gluts)? Logics capable of dealing with the latter of these two options will be of central importance to this work - the so-called paraconsistent logics. Taken to first- and higher-order settings, where we are not limited to discussion concerning complete sentences, but can also speak of objects and predicates, debate has naturally centred around the issue of *their* vagueness[2]. Might it be the case, for example, that the predicate "... is tall" is vague in a paraconsistent sense? There are certain individuals, no doubt, of whom we would say are (clearly) tall, and others of whom we would say

[1] Aristotle, who, it has been said created logic ex nihilo ([23], p.206), also considered future contingents in chapter IX of De Interpretatione, and whether statements like "There will be a sea battle tomorrow" could be neither true nor false. Such considerations would come to serve as Łukasiewicz's motivation for his logic Ł3.

[2] Note that there may be many different ways to understand phrases like "a vague object". The question of what it is to be a vague object (or, likewise, a vague predicate) is not uncontroversial. See, for example, [44] for some extensive thoughts on the matter. For better or worse, in this paper, we will consider the idea of vague object and vague property to be, in a sense, co-extensional. Where some predicate P both holds for an object a and not, i.e. $Pa \wedge \neg Pa$, then both the object and the predicate will be said to be vague.

are (clearly) not tall, but are there individuals of whom we might say "well, yes and no - they *are* tall and they aren't"? It at least seems plausible that this is the case. Tallness is not a peculiar predicate in this respect; We could make similar remarks concerning the predicate "...is large", or "...is young", to name just two. In short, vagueness appears to be prevalent in the world. Where talk of objects is permissible, it is no great leap to consider the idea of identity between objects. But within the context of some or other paraconsistent logic, it is not altogether clear how to proceed. How would allowing for objects that both have and do not have some property affect identity? In other words, how might we expect identity to behave, given paraconsistency?

Although this is one motivation for considering identity - that it is a puzzle as to how it might work where paraconsistency is around - this would not be much of one, if we were entirely content with how identity is treated in more familiar logics. That is, identity would appear to hold little interest, if, say, first-order logic with identity ($FOL_=$) was entirely problem-free. On the contrary, $FOL_=$ appears to face many problems of identity. To name but a few, we have sorites paradoxes; the Ship of Theseus; Church's (other) paradox; the problem of identity through time; the possibility of contingent identity; Evans' argument against vague identity; Geach's argument against absolute identity; the problem(s) of personal identity. In what follows, we will not attempt to propose solutions for all of these problems. Indeed, for most of them, the best we can offer is a starting point with which to base further research. Others will be entirely off limits to us, and for good reason. One problem that we *will* hope to deal with is a version of the sorites paradox. Later, we will see why accounting for this paradox can facilitate the treatment of other problems associated with change and identity. Consider the following thought experiment.

Suppose you are shown a wall upon which hang many sheets of paper - 256 of them to be exact - all in a horizontal line. Each sheet has been filled entirely with colour according to the RGB colour model. The first (the furthest to the left, say) has a colour value (a 3-tuple) of $\langle 255, 0, 0\rangle$, the second has a colour value $\langle 254, 0, 0\rangle$, and so on, until the last sheet - the 256th sheet - has value $\langle 0, 0, 0\rangle$. Suppose that you come to make 'two' observations: (i) the first sheet is red, and (ii) every adjacent sheet is identical. This leads you to make the following (sorites-like) argument to yourself:

$$\begin{array}{l} Ra_{255} \\ a_{255} = a_{254} \\ a_{254} = a_{253} \\ \dots \\ a_2 = a_1 \\ a_1 = a_0 \\ \hline Ra_0. \end{array}$$

For short, we will represent the argument as: $Ra_{255}, \forall n(a_n = a_{n-1}) \models Ra_0$, with the assumption that $0 \leq n \leq 255$. In words, it says that if the first sheet (a_{255}) is red, and any two sheets next to each other are identical, then you can correctly infer that the final sheet (a_0) is red too.[3]

Contrary to what the argument tells you, though, when you look at the final sheet (because, after all, it has value $\langle 0, 0, 0 \rangle$), it appears to you black, or in other words: most definitely *not* red. The paradox has thus been established: you have reasoned - validly, you suppose - towards a false conclusion from seemingly true premises.

1.1 The standard account, Leibniz's Law, and the Ship of Theseus

In both $FOL_=$ and second-order logic (SOL) the above argument is valid. In light of this, and also because they are both bivalent logics, we will refer to their approaches collectively as "the standard account". There are many features of the standard account that we might focus on to explain why the argument goes through. Here, we will look at two: the substitutivity of identity, and the transitivity of identity. The first of these principles says that whenever some predicate P holds of an object named a (i.e. Pa is true), and a is identical to b, then we can infer that Pb is true also. So, in our example, because Ra_{255} is true, and $a_{255} = a_{254}$, then we infer that Ra_{254} is true. But notice that this process can continue for as long as our identity

[3]It should be mentioned that implicit in the example is the fact that most people can't distinguish a single colour-value shift like the one described. This might suggest a favouring of a *de dicto* as opposed to *de re* reading of vagueness, but this is unintended. There may also be questions as to whether the sorites generally is better viewed as involving conditionals (- the more common reading, to be found in, for example, [35], [44], [38]) or identity statements (- the rarer view, for example, [26]), but this is not important for what we say. There may be questions as to whether this specific example should be considered in the conditional way, but we will come to this point later.

statements do. That is, from Ra_{254} and $a_{254} = a_{253}$, we infer Ra_{253}. From Ra_{253} and $a_{253} = a_{252}$, we infer Ra_{252}, and so on, until we reach Ra_1 and $a_1 = a_0$, and we conclude Ra_0. This is our troubling conclusion. The second of our principles says that if an identity holds between two named objects, and from the second named object to a third, then we can infer that identity holds between the first and the third. So, in our example, because $a_{255} = a_{254}$ and $a_{254} = a_{253}$, then we can infer that $a_{255} = a_{253}$. Similar to before, we can apply this many times to reach the conclusion that $a_{255} = a_0$. With one application of substitutivity, then, we reach our troubling conclusion again. In short, from Ra_{255} and $a_{255} = a_0$, it follows that Ra_0. Note that $a_{255} = a_0$ is a troubling conclusion in itself. It means that the (really) red sheet is the same as the (really) not red one. In fact, it is clear that the transitivity of identity will make it so that every sheet is identical to every other one. This seems to be a very troubling conclusion.

Given this situation, we might want to ask the following question: what rationale does the standard account have for finding the two principles we have mentioned correct? A plausible answer is that the standard account commits to Leibniz's law. That is, that objects are identical iff they have all the same properties, i.e. the conjunction of the identity of indiscernibles and the indiscernibility of identicals. Strictly speaking, of course, because $FOL_=$ does not allow for quantification over properties, it can only do this implicitly. For example, Leibniz's law often serves as the motivation for identity's introduction and elimination rules. In SOL, where it is fully expressible, it is the go-to definition for identity, usually taking the form of $a = b$ iff $\forall P(Pa \leftrightarrow Pb)$. Another way to put the definition is that two objects (a and b) are identical (one and the same) iff for every predicate, P, Pa will receive the same truth-valuation as Pb. Although these two definitions are usually equivalent, we will come to see that this need not be the case. Where we refer to Leibniz's Law in this paper, we will usually mean the first. Where we do not, that will be made clear. Although this version of identity is not without controversy[4], we will, in this paper, only be looking at logics which define identity in this way. This is perhaps surprising, given what we have said about the standard approach's connection to Leibniz's Law, and its failure to deal with the sorites. We will need to make one further qualification to the definition of Leibniz's Law we have given, namely that the properties we have in mind are in a sense simple. More specifically, our properties

[4] A historically famous counterexample is Black's symmetric universe, [5]. None of the paraconsistent approaches appear to help us deal with this problem, though there may be more standard ways of dealing with it. See, for example, [16].

will be extensional as opposed to intensional predicates. This will mean, for example, that we won't be considering modal properties. As to what sort of predicates we *will* consider, this will be made clear. For now, let us return to our version of the sorites paradox, the problem involving red sheets.

Though the particular example we gave involved the predicate redness and sheets of paper, the argument can be adapted, with suitable care, to cover any number of different predicates and objects, for example the predicate "...is a heap" and grains of sand[5], or number of hairs and heads. It is worth noting that in the early history of the sorites (stemming from Eubulides of Miletus - from whom it appears to have originated), it was framed and considered more of a puzzle than a paradox, the puzzle being specifically: where exactly is the cut-off point between, say, a red sheet of paper and a non-red one? In more recent times, the paradox approach, focusing instead on where the argument as a whole errs (with premise(s) and/or inference rule), has been far more popular. One notable exception that at least attempts to address both the paradox and the puzzle is the philosophical position of Epistemicism (for example, found in [44]), which claims that there *is* a sharp cut-off point, we just can't know where it is. Although Epistemicism can be praised for at least addressing the puzzle of the sorites, because it rejects vagueness at a fundamental level, we do not consider it a viable solution to the sorites.

We can use the version of the sorites we have given to make sense of other identity problems. Take for example, the idea of a ship undergoing change through time - the Ship of Theseus. Suppose that every so often, one of the old wooden planks of a ship is replaced with a new one. This process continues for a hundred years until every plank has been replaced. A plausi-

[5] The case of the heap might be objected to on the grounds that one could never mistake one grain, say, for two. In other words, while more standard sorites work because we need only demand that a single grain and two are identical with regards to the predicate "...is a heap", ours doesn't because we need a stronger sense of identity. While we might agree for the case given, it does not seem implausible that the argument could be changed to allow for this. Consider a descending sorites that starts with, say, 100 grains and ends with, say, 13 grains, and suppose that you were not able to stare at the heap for any great length of time. Given empirical work on 'subitizing' [19] this would now seem more plausible an example. This gives a sense of what we might mean by "suitable care" in adapting the argument. Similarly, for the red sheets example, we might want to say this is just an identity-in-terms-of-redness relation. After all, the sheets have different positions on the wall, so they can't be identical. Again, it appears that suitable changes could be made. This example seems to be easily translatable to a more abstract plane, where the sheets are atemporal aspatial objects.

ble reading of this is as a sorites-like transition: at each temporal stage the ship is identical to its previous self (because who's going to notice a single plank?) and yet the latest version of the ship is *entirely* different to the original. In this example, of course, our ship is analogous to the sheet of paper; the ship through time is the collection of sheets on the wall; and our individual planks are the individual colour-value changes. This should mean, then, that if we can have a solution to the version of the sorites paradox we have given, we can solve the problem of Theseus' Ship. Furthermore, in dealing with the sorites, we should also be able to deal with examples with many properties, for example, the one given by Priest [31] involving his motorbike. Over the course of time, this bike undergoes many changes: the tyres are replaced, the handlebars, and the seat, etc. The addition of further properties need not matter for us, whether they be one after another or concurrent through time. All that is crucial is that none of the changes are significant enough to cause a failure of identity between any stage and the one next to it.

It should now be mentioned that this is not the only possible reading of Theseus' Ship. Part of why it is a paradox, no doubt, lies in the idea that some crucial part of the ship never changes. Such thoughts naturally lead us towards personal identity. Are we to imagine ourselves as Theseus' Ship - never changing, yet always remaining the same? Such ponderings will not be possible in this paper, in light of what we have said about Leibniz's Law. In more Aristotelian terms, such thoughts involve properties that are both essential and accidental, and if we want to have identity hold in terms of all properties, then by saying the ship (or we) never change, we would be ignoring all the accidental properties that do. Our initial reading of the sorites only makes sense, then, if we limit our properties to extensional predicates, avoiding the mysterious thing-in-itself of intensional predicates. Another way in which the intensional reading of Theseus' Ship appears problematic is in light of Hobbes' addition to the problem, which is the idea that all the old bits from Theseus' Ship are assembled together to form a new, replica ship. Under the intensional view, which might define Theseus' Ship as, say, objects arranged in a Theseus Ship like way, we will now find there are two Theseus' Ships. Under the extensional view that we have taken, this second ship would, in fact, be the real Theseus' Ship, i.e. the original one, because it has exactly the same properties or material constituents as the original ship. There are many other issues we could discuss here involving the issue of material constitution, and whether the identity we are dealing with is an absolute one (as we have suggested) or just a relative one according to some narrow criterion, or predicate. With what we have, we are now in a position to more closely state the central question of this paper:

> Can we provide an account that deals with identity problems that also satisfies our natural intuitions regarding identity?

We have seen some example problems that we hope to deal with, but we have not yet fully addressed what counts as "our natural intuitions regarding identity". We can separate this into two parts. Firstly, we can think of how we should like identity to be defined, and secondly, we can think about how it should behave. To be sure, the behaviour of identity in any given logic will be wholly determined by how it is defined in that logic. We have already said that we want identity to be defined according to Leibniz's Law. But how identity behaves, then, will as much depend on the inner workings of the logic considered. It will thus make sense for us to isolate some specific behaviours that we think identity should have. To this end, we will focus on five sequents (which, strictly speaking, includes one meta-sequent) that together will form what will be called the "target features". They are: the substitution of identicals (SI); the transitivity of identity (TI); the non-absurdity of vague identity (NVI); the non-absurdity of contradictory objects (NCO); and finally, the transitivity of inference (ToI). Some initial remarks concerning each will be appropriate here.

1.2 The Target Features

We saw the substitutivity of identicals (SI) in the red sheet example earlier. In sequent form it can be represented as $Pa, a = b \models Pb$. If some property holds for some named object and this object is identical to 'another', then the property must hold of that 'second' object. Our use of quotation marks is, of course, indicative of the fact that, in an important sense, there is but one object, they are only named differently. We will see how paraconsistency might seem to put this into question somewhat, but this is the idea. It is important to distinguish (SI) from Leibniz's Law (as we have defined it). It is perfectly possible for two object-predicates to be equivalent for all predicates, and yet for (SI) to fail. Indeed, we will see an example of this in the next chapter. There have been numerous proposed counterexamples to (SI) in the literature, many focusing on what we might call naming predicates. To describe just two: Cicero = Tully, and "Cicero" has six letters, yet "Tully" does not have six letters. $a = b$, and "a" is the first letter of the alphabet, yet "b" is not [30]. Such examples appear to take the same form, and appear to be straightforwardly soluble in a Fregean manner, namely using the sense/reference distinction (i.e. the intensional/extensional distinction). In each case, it is the objects themselves (i.e. the referents) that are said to be identical (one and the same), and so it is the properties of the referents and not their senses that are said to be shared. (Again, it should be stressed that the idea is that there is only one referent (or extensional object).) There are further cases involving naming predicates that require more unpacking, the notable example being: Giorgione = Barbarelli; Giorgione was so-called because of his size; yet Barbarelli was not so-called because of his size. Though no explicit quotation has occurred here, a naming-predicate has been employed, which we will, by the reasons discussed before, disallow. That is, because we wish to take an extensional reading of the Ship of Theseus, we will only be focusing on extensional predicates. To further clarify, we might use Quinean terminology and say that the predicates, in that they will be extensional, will all be referentially transparent (as opposed to opaque). In effect, this says that wherever we wish to substitute some object for another into a sentence, we can only do so when that sentence doesn't involve predicates that won't allow this. (In a sense, it might seem trival that (SI) should hold then, but this needn't be the case. Again, in the next chapter.) We can use this idea of referential transparency to also deal with other non-extensional predicates. For example, intentional ones. These are predicates which in some way or another involve mental states. There are numerous examples of the failure of (SI) with such predicates. A historically famous example (which, interestingly, was also proposed by Eubulides) is the paradox of the hooded man. A man knows his brother who is standing in front of him. But he doesn't know who the person is because he's wearing a hood. It follows that (SI) does not hold for the property of being known. (See [29] for more details on this specific paradox.) This appears to share some commonality with Church's (other) paradox. Suppose an object has two names, a and b, and suppose that I know various things about a and b, but not that they are equal. $a = b$ iff they share *all* the same properties, though, so they can't be identical because they're not identical in my mind. There are other apparent counterexamples that involve different modal properties than epistemic ones (see, for example, [43] [22]). Given that we will only focus on extensional predicates, such examples will not be counterexamples of (SI) for us, and they will not be of concern in this paper. Where our properties are extensional, (SI) would appear to hold, hence its inclusion in our list of target features.

We have also seen the transitivity of identity (TI) before (again, in the red sheet sorites). We will represent it as $a = b, b = c \models a = c$. In mathematical settings, such a sequent would appear undeniable. Indeed, it is common for the transitivity of a relation, together with its reflexivity and symmetry to be the necessary and sufficient conditions of an equivalence relation, and identity is often defined to be the smallest equivalence relation. In more quotidian settings, the transitivity of identity appears equally intuitive. If you heard some historian saying, "George VI was the last Emperor of India, and the last Emperor of India was the first Head of the Commonwealth" you would be inclined to think that George VI was the first Head of the Commonwealth. Though this example might bring up other questions (are these names or descriptions? is 'was' 'is'?) the idea should be clear enough: where an identity holds between two objects and between the second and a third, identity will hold between the first and the third. Despite the apparent certainty of (TI), counterexamples to it have been proposed, most notably, perhaps, Prior's[6] amoeba example. Suppose at some time (t_0) we have an amoeba (a) which at some later time (t_0) divides to form two children (b and c). The children, being separate entities, exist at separate locations (l_b and l_c, respectively). b (say) must be identical to a, for if c were to cease to exist this would be obvious. If we were to ask at t_1, "where is a now?", the answer would be: "there, of course!" (pointing at b). And that the identity between any two things relies on something else is absurd. This is how the argument goes, at least. If we accept this for b, then we can make the similar argument for c, and so $a = b$ and $a = c$. But $b \neq c$ because they are clearly separate: they occupy separate positions - one is at l_b and the other at l_c. It follows that the transitivity of identity has failed: we have (in this case), $a = b, a = c \not\models b = c$. There is more to say concerning this (for example, what about property l_a?), but for now, the point is fairly clear: there may well be cases where the identity of transitivity might fail. Recent empirical research suggests that it fails for icebergs, for example [37]. We will see how our logics can deal with this example, but despite what we have said about possible counterexamples, given the intuitiveness of (TI), we will view it that (TI) should be the case.

We will represent the non-absurdity of vague identity (NVI) as $a = b, a \neq b \not\models \bot$. $\bot$ here will be used as shorthand for the idea that anything will follow from the premises, i.e. that $\Gamma \models \bot$ iff $\Gamma \models \phi$ for any arbitrary ϕ (- often referred to as "explosion"). Note that in none of the logics we consider will $\bot$ have any place in the language of those logics as a constant, nor will it be the

[6]The example, to be sure, relies as much on Priest as it does Prior. Prior made no mention of amoebas, but rather only referred to hypothetical individuals: "Let us suppose that the single individual x has become two individuals y and z." ([33], p.83) It might thus be better called "Prior-Priest's amoeba example". It is also worth noting that Prior's motivation for it was as a counterexample to what we have called (SI), and specifically as a counterexample that didn't rely on descriptions. It will also hold some importance for us for this reason as we will see.

case that, say, $\phi \rightarrow \bot$ is equivalent to $\neg\phi$ in any of the logics we consider. Informally, we might want to understand $\bot$ as implying that a serious problem has occurred, something much worse than 'mere' paraconsistency. We may in places refer to this as a contradiction, but as we will shortly see, our usage of this term is not consistent. It should also be noted that $a = b, a \neq b$ is not the same idea of vague identity that is, for example, used by Evans [13]. There the idea is of indeterminate identity, where indeterminacy is a sentential operator; Here it is an example of a fairly common set of premises to see in a paraconsistent logic. We have already seen - in Prior's amoeba example - some evidence of why we might want (NVI) to be the case. Observe that it seems to be the case that $a = b$ and $a \neq b$ because $\neg l_b a \wedge l_b b$ (because a has all the properties of c). If we took this to be problematic, then it doesn't seem like we would be able to deal with the amoeba example. Despite this, the idea that objects can be indistinguishable and yet distinct might seem plainly absurd. The special case where a and b are symbolically indistinct, i.e. $a = a, a \neq a \not\models \bot$ appears especially troubling. Some treatment of this view will be considered throughout, but also in the conclusion.

On its own, (NCO) - represented as the sequent $Pa, a = b, \neg Pb \not\models \bot$ - can also be considered a fairly common principle in any paraconsistent logic. Its relevance to the sorites was implicit when we discussed the red sheet example earlier. In a sense, we can think of it as being a miniature sorites in one, incorporating its main features. Despite the fact that all our logics will be able to deal with the sorites, we will see what use (NCO) can have in further distinguishing them. In the fourth chapter, (NCO) will play an important role in dealing with the sorites, which, roughly speaking, will be a result of how affirming (SI) together with a variation of (NCO) has dramatic consequences for the inner workings of the logic of that chapter.

The last of our target features - the transitivity of inference (TOI) - is neither a sequent, nor does it explicitly refer to identity at all. We will represent it here as the meta-sequent $\phi \models \psi, \psi \models \chi \Rightarrow \phi \models \chi$, which, in words, means that where we have an inference from one sentence to another, and also an inference from that second to a third, it follows that from the first sentence we can infer the third. We have chosen this formulation of inferential transitivity, as opposed to, say, the more general $\Gamma \models \Gamma', \Gamma' \models \Gamma'' \Rightarrow \Gamma \models \Gamma''$, where $\Gamma, \Gamma', \Gamma''$ are all sets, because all the logics we will be considering will be single-conclusion logics. It might also need mentioning that because of how our logics are constructed, a more technically correct representation of (ToI) would be $\{\phi\} \models \psi, \{\psi\} \models \chi \Rightarrow \{\phi\} \models \chi$. This will be implicit whenever we use it. We won't, at this stage, dwell upon (ToI)'s importance, save to say that for the logic we consider in the third chapter it will be crucial to our resolution of the sorites. Related to this, it can also serve as a feature with which to distinguish the logic of the third chapter from the standard account.

At this early stage, we are in a position to mention how the standard approach deals with our target features. While (SI), (TI), and (ToI) hold, (NCO) and (VI) both fail. It can be said, then, that the standard account fails on two separate fronts: it is incapable of dealing adequately with identity problems (specifically, it is unable to deal with our version of the sorites paradox and hence Theseus' Ship), and it fails to satisfy our intuitions in terms of the target features. In what follows, we will see what hope paraconsistent logics might have on both these counts.

Chapter 2

Logic of Paradox - Embracing Inconsistency

The logic of Paradox (LP) is probably the most well-known paraconsistent logic. Although part of the semantics for it were initially considered by Asenjo ([3]), it is now almost exclusively associated with Graham Priest, who has argued extensively for its employment over classical logic in a wide variety of different contexts (for example, the liar paradox in [25], and even pure mathematics, [28]). The present chapter will focus, in the main, on a version of LP found in [31]. The main idea there was that we can use some of the properties of propositional LP - specifically, the non-standard way the conditional behaves - to define identity in the Leibnizian way using the bi-conditional within second-order LP, and then use identity to solve problems. Broadly speaking, this will in fact be how we proceed for all the paraconsistent logics in this paper. The procedure can be thought of as a three-step process: (i) Take some paraconsistent logic extrapolated to second-order; (ii) Define identity in the Leibnizian way; (ii) Use identity as defined to solve previously intractable identity problems. For now, we will look in detail at how the process can be carried out with LP, starting with our second-order formulation.

2.1 Formalising LP

Our language will contain individual constants $(a, b, c, ...)$ and variables $(x, y, z, ...)$, monadic predicate constants $(A, B, C, ...)$ and variables $(X, Y, Z, ...)$, connectives $(\wedge, \vee, \neg)$, and finally the first and second-order quantifiers $\forall, \exists$. The connectives $\rightarrow$ and $\leftrightarrow$ will be considered as derived, namely that $\phi \rightarrow \psi := \neg\phi \vee \psi$

and $\phi \leftrightarrow \psi := (\phi \to \psi) \wedge (\psi \to \phi)$. Function symbols will be avoided, as will formulas involving free variables.

An interpretation for the language, $\mathcal{I}$, is a triple $\langle D_1, D_2, \theta \rangle$, where D_1 is the non-empty domain of first-order quantification, and D_2 is a set of pairs of the form $\langle A^+, A^- \rangle$, where $A^+ \cup A^- = D_1$. Intuitively, A^+ (which we'll call the extension of A) and A^- (its co-extension) contain the objects of D_1 for which the property A holds, and does not hold, respectively. In classical logic, we would only need to specify the extension - any object not in the extension would be in its co-extension, and no object could possibly be in both. Given paraconsistency, we specify the co-extension, meaning we can have objects in this, or the extension, or both. Note that because $A^+ \cup A^- = D_1$, it is not possible for an object to be in neither. (In the next chapter the consequences of dropping this restriction will be briefly mentioned.) We will also make the restriction that for every $A \subseteq D_1$, there is a $B \subseteq D_1$ such that $\langle A, B \rangle \in D_2$, but make no further claims about how extensive D_2 is. (For example, if $D_1 = \{d_1\}$, it needn't be the case that both $\langle \{d_1\}, \{\emptyset\} \rangle \in D_2$, and $\langle \{d_1\}, \{d_1\} \rangle \in D_2$, but our restriction does demand that at least one of them is in D_2. Intutively, the semantics employed is closer to full semantics (as opposed to Henkin semantics) in that the union of the predicate extensions ($\bigcup_{A \in D_2} A^+$) is equal to the power set of D_1. The union of the predicate co-extensions need not be though.)

θ will assign an element of D_1 to every individual constant, and an element of D_2 to every predicate constant. If A is a predicate, we will write $\theta(A)$ as $\langle \theta^+(A), \theta^-(A) \rangle$, and we define $\theta(Aa) = \theta(A)(\theta(a))$. Finally, an evaluation v assigns to every formula one of $\{0\}, \{1\}, \{0, 1\}$ (intuitively, strictly false, strictly true, borderline true, respectively) according to the following rules:

$$0 \in v(Aa) \text{ iff } \theta(a) \in \theta^-(A)$$

$$1 \in v(Aa) \text{ iff } \theta(a) \in \theta^+(A)$$

$$0 \in v(\neg\phi) \text{ iff } 1 \in v(\phi)$$

$$1 \in v(\neg\phi) \text{ iff } 0 \in v(\phi)$$

$$0 \in v(\phi \wedge \psi) \text{ iff } 0 \in v(\phi) \text{ or } 0 \in v(\psi)$$

$$1 \in v(\phi \wedge \psi) \text{ iff } 1 \in v(\phi) \text{ and } 1 \in v(\psi)$$

$$0 \in v(\phi \vee \psi) \text{ iff } 0 \in v(\phi) \text{ and } 0 \in v(\psi)$$

$$1 \in v(\phi \vee \psi) \text{ iff } 1 \in v(\phi) \text{ or } 1 \in v(\psi)$$

For example, suppose we have some predicate A and an object a, and our interpretation $\mathcal{I}$ is such that $\theta(a) \in D_1$ and $\theta(A) = \langle \{a\}, \{a\} \rangle \in D_2$. It follows that, for this interpretation, $v(Aa) = \{0, 1\}$ as $\theta(a) \in \theta^+(A)$ and $\theta(a) \in \theta^-(A)$. Intuitively, A is thus a vague predicate - it can be said to both hold for the object a and not hold for it. As might be expected then, the sentence $Aa \wedge \neg Aa$ will also be true, that is, $1 \in v(Aa \wedge \neg Aa)$ under $\mathcal{I}$. More precisely though, $v(Aa \wedge \neg Aa) = \{0, 1\}$, as one of the conjuncts (in this case, both) will have value 0. In short, $Aa \wedge \neg Aa$ will be borderline true under $\mathcal{I}$. It can be shown that all the familiar features of propositional LP (- and first-order, for that matter -) will be carried over to the second-order version of LP we are considering; That is, we can safely say that this version of LP is an *extrapolation* into second-order. Before we come to define quantification, we will look at two propositional examples. Starting with the conditional, we said before that $\phi \rightarrow \psi$ would be understood as short-hand for $\neg\phi \vee \psi$. This means that: $v(\phi \rightarrow \psi) = \{1\}$ iff $v(\phi) = \{0\}$ or $v(\psi) = \{1\}$; $v(\phi \rightarrow \psi) = \{0\}$ iff $v(\phi) = \{1\}$ and $v(\psi) = \{0\}$; and finally that $v(\phi \rightarrow \psi)$ in every other case, i.e. where one of (or both) $v(\phi) = \{0, 1\}$ and $v(\psi) = \{0, 1\}$. In words then, $\phi \rightarrow \psi$ is strictly true (i.e. $v(\phi \rightarrow \psi) = \{0, 1\}$) iff ϕ is strictly false or ψ is strictly true; $\phi \rightarrow \psi$ is strictly false iff ϕ is strictly true and ψ is strictly false; and $\phi \rightarrow \psi$ is borderline true in every other case. All this is in agreement with the perhaps more standard truth-table approach.[1] Following from these results, we have it that $\phi \leftrightarrow \psi$ is strictly true iff both ϕ and ψ are strictly true, or both are strictly false; $\phi \leftrightarrow \psi$ is strictly false iff ϕ is strictly true and ψ is strictly false, or *vice versa*; and finally that $\phi \leftrightarrow \psi$ is borderline true where one (or both) ϕ and ψ are borderline true. Intutively, for $\phi \leftrightarrow \psi$ to have some truth (i.e. to be at least tolerantly true), ϕ and ψ must share some common truth-value, that is, 0 or 1.

For the quantifiers, we will assume that the language is expanded if necessary to give each member of D_1 and D_2 a name. If $d \in D_1$, we will write its name as $\mathbf{d}$ and if $A \in D_2$, we'll call it $\mathbf{A}$.

$$0 \in v(\exists x \phi(x)) \text{ iff for all } d \in D_1, 0 \in v(\phi(\mathbf{d}))$$

$$1 \in v(\exists x \phi(x)) \text{ iff for some } d \in D_1, 1 \in v(\phi(\mathbf{d}))$$

$$0 \in v(\forall x \phi(x)) \text{ iff for some } d \in D_1, 0 \in v(\phi(\mathbf{d}))$$

$$1 \in v(\forall x \phi(x)) \text{ iff for all } d \in D_1, 1 \in v(\phi(\mathbf{d}))$$

$$0 \in v(\exists X \phi(X)) \text{ iff for all } A \in D_2, 0 \in v(\phi(\mathbf{A}))$$

[1] See, for example, p277 in [25]. Note, the terminology used there is somewhat different from the one we use in this paper.

$$1 \in v(\exists X\phi(X)) \text{ iff for some } A \in D_2, 1 \in v(\phi(\mathbf{A}))$$

$$0 \in v(\forall X\phi(X)) \text{ iff for some } A \in D_2, 0 \in v(\phi(\mathbf{A}))$$

$$1 \in v(\forall X\phi(X)) \text{ iff for all } A \in D_2, 1 \in v(\phi(\mathbf{A}))$$

For example, suppose we have an interpretation, $\mathcal{I}$, such that $d_1 \in D_1$, $A = \langle d_1, \emptyset \rangle \in D_2$, . It will follow that the sentence Pa will be strictly true, that is, $v(A\mathbf{d}) = \{1\}$, because $\theta(d) \in \theta^+(A)$, and $\theta(d) \notin \theta^-(A)$. It will also follow then that $v(\exists x(Ax)) = \{1\}$ because $d \in D_1$ such that $1 \in v(A\mathbf{d})$, and $0 \notin v(A\mathbf{d})$. Suppose also though that for $\mathcal{I}$ there is some other predicate $B \in D_2$ such that $\theta(B) = \langle \{d\}, \{d\} \rangle$. It will follow that Ba is true and false, i.e. borderline true. It will follow that $\neg\forall X(Xa \wedge \neg Xa)$, because $1 \in v(Ba)$ and $0 \in v(Ba)$. Note also, though, that it will still be the case that $\forall X(Xa \leftrightarrow Xa)$ because for every predicate X, Xa still shares a truth-valuation with itself.

2.2 *LP* Validity, Consequence, and the Biconditional

For validity, we will say that an interpretation is a model of ϕ iff $1 \in v(\phi)$. Where Σ is a set of formulas, $\mathcal{I}$ is a model of Σ iff it is a model of every member. Finally, $\Sigma \models_{LP} \phi$ will hold iff every model of Σ is a model of ϕ. (Note that we will usually drop set brackets from premises.)

Intuitively, entailment in *LP* preserves truth, though not necessarily strict truth. (Put another way, *LP* preserves non-strictly false interpretations.) A famous sequent in *LP* that shows this is the failure of *modus ponens*: $\phi, \phi \to \psi \not\models_{LP} \psi$. Suppose we have an interpretation, $\mathcal{I}$, for which $v(\phi) = \{0, 1\}$ and $v(\psi) = \{0\}$. $\mathcal{I}$ will be a model for ϕ because $1 \in v(\phi)$. It will also be the case that $1 \in v(\phi \to \psi)$ because $0 \in v(\phi)$ (this was shown previously). It follows then that $\mathcal{I}$ is a model for both ϕ and $\phi \to \psi$, hence it is a model for the set containing them. But, by our assumption, $1 \notin v(\psi)$, hence $\mathcal{I}$ is not a model for ψ. It follows that there is a model for the premises that is not a model for the conclusion, and so *modus ponens* fails. Interestingly though, $\models_{LP} (\phi \wedge (\phi \to \psi)) \to \psi$ will hold. This is because under any interpretation, $(\phi \wedge (\phi \to \psi)) \to \psi$ will at least be borderline true. For example, if we take our interpretation $\mathcal{I}$ from above, $\mathcal{I}$ will make $v(\phi \wedge (\phi \to \psi)) = \{0, 1\}$, and, as said before, if 0 is in the antecedent of a conditional, then 1 will be in the conditional. It will be the case then that $v((\phi \wedge (\phi \rightarrow \psi)) \rightarrow \psi) = \{0, 1, \}$ under $\mathcal{I}$. In fact, all propositional and first-order LP validities will be classical ones, and *vice versa*. The same could be said for second-order LP and second-order logic, given our second-order logic is defined in the right way. (See [31] p.7).

As we proceed, the following sequents involving the biconditional will be of particularly importance to us:

$$\models_{LP} \phi \leftrightarrow \phi.$$

$$\phi \leftrightarrow \psi \models_{LP} \psi \leftrightarrow \phi$$

$$\phi, \psi \models_{LP} \phi \leftrightarrow \phi.$$

$$\neg\phi, \neg\psi \models_{LP} \phi \leftrightarrow \psi$$

$$\phi, \neg\psi \models_{LP} \neg(\phi \leftrightarrow \psi)$$

$$\phi, \neg\phi \models_{LP} \phi \leftrightarrow \psi$$

$$\phi, \psi \models_{LP} \neg\phi \leftrightarrow \neg\psi$$

$$\phi \leftrightarrow \psi, \psi \leftrightarrow \chi \models_{LP} \neg(\phi \leftrightarrow \chi)$$

The first of these implies that the biconditional is reflexive; The second, that it is symmetric; The next two suggest that if two sentences have the same truth-value, the biconditional will hold between them; The fifth implies that if two sentences have different truth-values then the biconditional will not hold; The sixth, that where we have a borderline true sentence (ϕ), any sentence (ψ) will be equivalent to it (this follows from what we were saying before about the biconditional needing sentences to share a common truth-value - 0 or 1); The seventh shows that the contrapositive holds for the biconditional; Finally, the eighth sequent implies that the biconditional is not transitive. Together with the sixth result, these two are the only non-classical results in the sequents listed above. The non-transitivity of the biconditional, we will see, does most of the heavy lifting when it comes to the examples. Consider some interpretation, $\mathcal{I}$, where $v(\phi) = \{1\}$, $v(\psi) = \{0, 1\}$, and $v(\chi) = \{0\}$, i.e. where ϕ is strictly true, ψ is borderline true, and χ is strictly false. As shown previously, the biconditional will hold between two sentences wherever they share a common truth-value. It follows then that I is a model for each $\phi \leftrightarrow \psi$ and $\psi \leftrightarrow \chi$ (and so for the set containing them). In short, our premises hold. But clearly $\mathcal{I}$ is not a model for $\phi \leftrightarrow \chi$, because $v(\phi) = \{0\}$ and $v(\chi) = \{1\}$, i.e. they do not share a common truth-value. Hence our conclusion is false, and the biconditional can be said to not be transitive. Of course, there will be other interpretations (for example, where $v(\phi) = v(\psi) = v(\chi)$) where transitivity does hold for the biconditional, so it will be the case that the biconditional is non-transitive rather than intransitive. We will omit the proofs for the other sequents. They can be carried out in a similar manner.

2.3 *LP* and the Identity Problems

We are now in a position to define identity in the Leibnizian way, i.e. that two objects are identical iff they have the same properties:

$$a = b \text{ iff } \forall P(Pa \leftrightarrow Pb).$$

In so far as we have defined it, then, identity could be said to be standard. However, because of the underlying non-bivalent nature of the logic out of which it has been formed, it ends up behaving non-standardly. Roughly speaking, we can think of the properties of the biconditional as carrying over to identity. For example, identity will be reflexive and symmetric in LP. Just as with the biconditional however, this should not be misunderstood as meaning that under every possible interpretation $a = a$, say, will be *strictly* true. In fact, for $v(a = a) = \{1\}$ to be the case, we would need something much stronger, namely that $\neg\exists X(Xa \wedge \neg Xa)$, i.e. that a holds no vague predicates, or equivalently (see introduction) that a is not a vague object. For suppose we have some interpretation, $\mathcal{I}$, where $d \in D_1$, and $A = \langle\{d\}, \{d\}\rangle \in D_2$. It will be the case that $1 \in v(a = a)$ because they share a common truth-value, but also that $1 \in v(a \neq a)$, because they don't. One way of putting this is that $a = a$ is borderline true under $\mathcal{I}$, but perhaps a more telling way of putting it is that, under $\mathcal{I}$, $a = a \wedge a \neq a$ is (borderline) true. Indeed it is the case that in LP vague objects are both self-identical and not self-identical. (We could make similar remarks regarding the symmetry of identity in LP.) Although this simultaneous success and failure of self-identity is perhaps unproblematic for the logician, or one who is interested only in the mechanics of LP, (they can, after all, merely cite it as a characteristic of vague objects), for the philosopher, or one who wants to explain what is going on intuitively, this might be appear to be a problem. What does it mean for an object to be identical to itself and yet be distinct? Does it provide any insight into the nature of vague objects? We will return to this point in the concluding chapter of this paper. For now, we will look at some more concrete examples of the behaviour of identity, with an eye toward dealing with the more full-blown problems we were considering in the introduction. Lets see some real-world examples, with a view to getting towards the examples we saw in the introduction.

To start with, lets consider an object that undergoes change. Suppose we start out with an object, a that, at some time, is consistent. That is, for every predicate a has, a has only that predicate, and does not have also its negation. Let's focus on one of these predicates, A for which, by assumption Aa holds. (The example runs equally well, mutatis mutandis, if we let $\neg Aa$ hold initially.) Suppose though that at some later time, our object comes to hold the predicate $\neg A$ in addition to all the others it previously did. Call this "new" object b. It will then be the case that $\forall X(Xa \wedge Xb)$, and yet $\exists X(Xa \wedge \neg Xb)$. It follows that $\forall X(Xa \leftrightarrow Xb)$, and so $a = b$, and also $\exists X(Xa \leftrightarrow \neg Xb)$, and so $\neg(a = b)$. In other words, a will at once be identical to b and also distinct from it. It follows that (NVI) holds for LP. This example also shows us how (NCO) fails for LP. We can have identical objects, but that have contradictory properties. In this case, we have Aa, $a = b$, and $\neg Ab$ all being true (at least borderline true), but because of the paraconsistent nature of LP, we can't conclude everything.

Let us suppose now that our object comes to fully lose that property A. In a similar vein, we hypothesize a third object, c, and stipulate that $\neg Ac$ is the case, but not Ac, all other predicates remaining the same. It will now follow that $b = c$, because the two objects share the property $\neg A$ (and all other predicates remained constant), but $a = c$ will not be the case because it is not the case that $\forall X(Xa \leftrightarrow Xc)$ (Aa is strictly true, while Ac is strictly false, hence they share no common-truth value). As we changed nothing involving a and b, $a = b$ will still hold. Hence, we will have it that $a = b$, $b = c$, and yet $\neg(a = c)$. In other words, the transitivity of identity has failed. To see this in more detail, lets consider interpretations again. Suppose $\mathcal{I}$ is an interpretation whereby $D_1 = \{d_1, d_2, d_3\}$, $\theta(a) = d_1, \theta(b) = d_2, \theta(c) = d_3$, $A = \langle\{d_1, d_2\}, \{d_2, d_3\}\rangle \in D_2$, and for all other $B \in D_2$, $B^- = D_1$. In words, our first-order domain has three objects (d_1, d_2, d_3) that are assigned our constants (a, b, c, respectively) by the function θ; our second-order domain contains a predicate A which has d_1 and d_2 in its extension, and d_2 and d_3 in its co-extension; by assumption, this is the only predicate that a, b, c have - d_1, d_2, d_3 are only in the co-extension of every predicate other than A. It will follow that for $\mathcal{I}$, $\mathbf{A}a$ will be strictly true (because $\theta(a) \in A^+$ and $\theta(a) \notin A^-$), $\mathbf{A}b$ will be borderline true (because $\theta(b) \in A^+$ and $\theta(b) \in A^-$), and finally $\mathbf{A}c$ will be strictly false (because $\theta(c) \in A^-$). For every other predicate B, $\mathbf{B}a$, $\mathbf{B}b$, and $\mathbf{B}c$ will all be strictly false. From our rules concerning the biconditional, it will thus follow that, under $\mathcal{I}$, $1 \in v(\forall X(Xa \leftrightarrow Xb)$, $1 \in v(\forall X(Xb \leftrightarrow Xc)$, and $1 \notin v(\forall X(Xa \leftrightarrow Xc))$. In other words, $1 \in v(a = b), 1 \in v(b = c)$, and

$1 \notin (a = c)$ under $\mathcal{I}$. To show that $a = b, b = c \not\models_{LP} a = c$ we need only an interpretation that is a model for $a = b$ and $a = c$, but not for $a = c$. I is such an interpretation. Therefore, identity is not transitive in LP: (TI) fails.

The above example will allow us to make sense of Prior's amoeba example. In the introduction we described a situation in which an amoeba existed at time t_0 which then divided into two 'new' amoebas at t_1. The resulting amoebas b and c, existing now as clearly separate entities, would have distinct locations - l_b and l_c, respectively. We can let these locations act much like the predicate A above. Amoebas b and c can then be consistent in terms of these predicates. That is, $l_b b$, $\neg l_c b$, $\neg l_b c$, and $l_c c$ will all be strictly true. We noted that an amoeba parent and the child that results from its splitting could be said to be identical. We said that this would be obvious if one of the two offspring were to cease to be, and that it wouldn't make sense for an identity statement to depend upon the existence or non-existence of some other entity. In this example, a can play the role b did in the example above, acting, in a sense, as the bridge between the two other amoebas. In other words, $l_b a$, $\neg l_b a$, $l_c a$, $\neg l_c a$ will all be the case, and a will be vague. If we assume these to be the only relevant properties, then $a = b$ and $a = c$, but it will not be the case that $b = c$, because b would not share any predicate with c. Put another way, $l_b b$ and $l_b c$ would share no common valuation, and nor would $l_c b$ and $l_c c$, so $\neg\exists X(Xb \leftrightarrow Xc)$, hence $b \neq c$. Given what we have said then, LP does a rather good job at making sense of Prior's amoeba example. The example can serve to show why (SI) fails in LP. For consider the property l_b, i.e. the property of having b's location. In that a and b are identical, a is said to have this property. But a is also identical to c, and yet c does not have this location. In other words, $l_b a, a = c \not\models_{LP} l_b c$, and (SI) fails in LP.
We now have all we need to make sense of the version of sorites paradox in the introduction, and so Theseus' Ship. Recall our argument:

$$\begin{array}{l} Ra_{255} \\ a_{255} = a_{254} \\ a_{254} = a_{253} \\ \ldots \\ a_2 = a_1 \\ a_1 = a_0 \\ \hline Ra_0. \end{array}$$

Given the at we have seen both (TI) and (SI) fail for LP, we might suspect, from what was said in the introduction, that this argument is not valid in LP. This is indeed the case. For it is not possible even to conclude Ra_254, because a_254 might, after all, simply be borderline true. This means that we definitely can't conclude Ra_0. For identity, we have a similar result. We can't even be sure that $a_{255} = a_{253}$, for example, let alone that every element in the series is identical. This does not mean that some interpretations won't allow for this to be the case; To see whether the argument as a whole is valid, we are of course only speaking of all possible interpretations. A plausible model for the above example would be one which had the first certain number of objects all consistent and strictly true; the next certain amount borderline; and the final consistent but strictly false. The result would be that those first objects would all be identical, and so would the last number, but the middle ones would be both identical and not. The presence of borderline cases would mean that the predicate would not bleed across the border. As to precisely what model to use, this would be an empirical matter, based on experiments with many subjects, but the point is that it is at least plausible that such research could be done. Where we want to adapt the sorites to fit other phenomena, such as the Ship of Theseus, or personal identity (with extensional predicates), the crucial feature of the sorites would remain in place. All that is important is that the change is gradual in the sense that there are no sharp cut-off points, and LP allows this.

From what has been said, LP is capable of at least plausibly dealing with identity problems, and in summary of its treatment of the target features, (SI) and (TI) fail in LP, while $(NVI), (NCO)$, and (ToI) all hold.

Chapter 3

Strict-Tolerant Logic - Saving Consistency

In the previous chapter we saw how paraconsistency could help us to to deal with some traditional identity problems, problems that were not resolvable through what we called the standard account. Loosely speaking, we can put the successes of LP in this regard down to its weakening of various classical laws. Specifically, we saw that the failure of (TI) in LP played an important role in our reading of the sorites and Theseus' Ship paradoxes, and we saw that this was, in a sense, a product of the non-standard way the biconditional behaves in LP. We also saw what benefits paraconsistency in LP had with respect to other identity problems, for example with Prior's amoeba example. We likewise feel that its treatment of (NVI) and (NCO) as holding is correct. Although for these reasons it might be said that LP should be commended, given what was said in the introduction regarding our target features, we do not think LP adequately captures our natural intuitions regarding identity. In particular, we saw that (SI) failed for LP, which we took to be, given certain qualifications, a perfectly reasonable principle. In this chapter, we will see what hope there might be for holding on to the more consistent target features of (SI) and (TI), while still being able to deal with identity problems. In short, this chapter will answer the question of what other possibilities might there be for dealing with identity problems that do not resort to denying (SI) and (TI).

It is important to note that we need not be confined to the notion of model taken by LP. That there are other options we may take, though, is somewhat obscured by the notation we have seen. For this reason, we will provide another, one that can be shown to be equivalent to it. Our usage will be the same as in [10], which is the most important paper for this chapter as a whole. Using this new notation will also bring things more in to line with respect to the field of many-valued logics generally. Before we do so, though, it is perhaps worth mentioning why we might, for philosophical reasons, prefer the Priest notation to the one we will give in this chapter. This is in the sense that in choosing a three- (as we will do), or generally, an n-valued description of situations appears us to commit to more truth-values than we might want to. Whether this commitment is purely a linguistic one, or is also ontological in some sense is debatable, but in either case, positing values over and above the traditional two might not be something that we want to do. Of course, formally, our values will behave as symbols do on the page, devoid of meaning, so to speak, so it is perhaps only a matter for philosophers. Priest's own position is that LP specifically is not a logic that has any more truth-values than the familiar two. In his words, "[...] there are, in fact, only two truth values, *true* and false. It is just that sentences may have various combinations of these."[1]. With other logics, for example, Bochvar's three-valued logic, the 'third' value is simply an absence of truth-value (see [41]). From this perspective, the name *Many*-Valued Logics could well be thought of as a misnomer. Priest's reason for his own view could be that it lessens the radicalness of LP. While LP is of course radical in that it allows sentences to have more than one truth-value, it goes no further: it does not go so far as to change the fundamental nature of what references (in a Fregean sense) sentences can have. Whether Priest's position is the right one or not, we will remain neutral on the matter. But it is at least worth pointing out that the position taken by Priest regarding LP appears to be equally applicable to any of the three main logics we consider in this paper. For some further thoughts on this issue, see, for example, [17] pp.213-215. We will now set out ST formally, making use of a Strong-Kleene valuation schema.

3.1 Formalising ST

Assume the (second-order) language of the previous chapter, i.e. we will have individual names/variables, predicate names/variables, and the connectives $\neg, \wedge, \vee$, and the quantifiers $\forall, \exists$. $\phi \rightarrow \psi$ will again be a notational shorthand for $\neg\phi \vee \psi$, and $\phi \leftrightarrow \psi$ will be short for $(\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi)$. Once more we will avoid polyadic predicates (i.e. relations) and function symbols.

An $MV_2 - model$ is a structure $\langle D_1, D_2, \mathbb{I}\rangle$ such that:

[1] [30] p.7

· D_1 is a non-empty domain of quantification.

· D_2 is a set of functions in $\{0, 1/2, 1\}^{D_1}$.

· $\mathbb{I}$ is an interpretation function that:

- For an individual name or variable a, $\mathbb{I}(a) \in D_1$.
- For a predicate name or variable P, $\mathbb{I} \in \{0, 1/2, 1\}^{D_1}$.
- For an atomic formula Pa, $\mathbb{I}(Pa) = \mathbb{I}(P)\mathbb{I}(a)$.
- $\mathbb{I}(\neg\phi) = 1 - \mathbb{I}(\phi)$.
- $\mathbb{I}(\phi \wedge \psi) = min(\mathbb{I}(\phi), \mathbb{I}(\phi))$.
- $\mathbb{I}(\phi \vee \psi) = max(\mathbb{I}(\phi), \mathbb{I}(\phi))$.
- $\mathbb{I}(\forall x A) = min(\{\mathbb{I}'(A) : \mathbb{I}'$ is an x-variant of $\mathbb{I}\})$.
- $\mathbb{I}(\exists x A) = max(\{\mathbb{I}'(A) : \mathbb{I}'$ is an x-variant of $\mathbb{I}\})$.
- $\mathbb{I}(\forall X A) = min(\{\mathbb{I}'(A) : \mathbb{I}'$ is an X-variant of $\mathbb{I}\})$.
- $\mathbb{I}(\exists X A) = max(\{\mathbb{I}'(A) : \mathbb{I}'$ is an X-variant of $\mathbb{I}\})$.

We can also put the constraint we saw before concerning relevant properties. Here, this will be that for each $A \subseteq D_1$ there is an $f \in D_2$ such that for each $a \in A$, $f(a) > 0$. The resulting account is equivalent to the one we saw in the last chapter. Instead of speaking of extensions and co-extensions of predicates as we did previously, though, we now speak of predicates as being made of functions that map objects in D_1 to the truth-values $0, 1/2, 1$. For example, if want to express there being a vague predicate in a model M, $P \in D_2$ - a predicate for which we want the sentence $Pa \wedge \neg Pa$ to hold, say - we will make it so that P contains a function that assigns the value $1/2$ to $a \in D_1$. It will now be the case that, under M, $\mathbb{I}(Pa \wedge \neg Pa) = 1/2$ will be the case because $\mathbb{I}(Pa) = 1/2$, $\mathbb{I}(\neg Pa) = 1 - 1/2$, and $\mathbb{I}(Pa \wedge \neg Pa) = min(1/2, 1/2) = 1/2$. We will define identity in the same way as before, i.e. $a = b$ iff $\forall P(Pa \leftrightarrow Pb)$. (There is no difference in definition between LP identity and ST identity.)

LP consequence can be re-expressed as:

$\Gamma \models_{LP} \phi$ iff there is no MV_2-model for which $\mathbb{I}(\gamma) > 1$, for every $\gamma \in \Gamma$ and $\mathbb{I}(\phi) = 0$.

In the previous chapter, we said that LP consequence preserved tolerant truth. That is, LP inference made it so that whenever all the premises were not strictly false, the conclusion would not be strictly false either. It is clear that the above definition, in that it relies on interpretations being greater than 0 is in agreement with this. Given we have three values to play with, however, it is equally clear that this needn't be the only sort of consequence we could think of. For example, suppose we made it so that consequence relied instead on interpretations having value 1. That is, suppose we insisted on preservation of strict truth. The resulting logic would be Kleene's $K3$[2]. One insight in [**?**] was that in principle we needn't restrict ourselves to the same interpretation for premises and conclusions, that we could, in this sense, have our interpretations respect positionality, and that actually logics with so-called "mixed" consequence relations might be worth studying in their own right. The strict-tolerant logic (ST) is one such logic, treating premises like $K3$ and conclusions like LP. We can define it as:

$\Gamma \models_{ST} \phi$ iff there is no MV_2-model for which $\mathbb{I}(\gamma) = 1$, for every $\gamma \in \Gamma$ and $\mathbb{I}(\phi) = 0$.

Intuitively, we might want to understand ST as making the following demands on an acceptable inference: it must have premises that are true enough to make a sound argument, while not having a conclusion that is false enough to allow for a counterexample.

3.2 ST and the Identity Problems

We said previously that, with a classical vocabulary, LP validities were all classical validities and vice versa. Given that, with respect to premise-less arguments, ST and LP are indistinguishable - they both take a tolerant reading of sentences in the conclusion position - it follows then that ST validities will likewise totally coincide with classical ones. But ST takes this idea further though. In fact, *all* inferences in ST are classical inferences. To get a sense of this, consider the simple example of modus ponens, i.e. $\phi, \phi \to \psi \models \psi$. (Recall that this failed for LP - perhaps an unwelcome feature.) For ST, we only need to look at an interpretation $\mathbb{I}$ such that $\mathbb{I}(\phi) = \mathbb{I}(\phi \to \psi) = 1$. From how we defined the conditional, $\mathbb{I}(\phi \to \psi)$ will be equal to whatever is the highest value of $\mathbb{I}(\neg\phi)$ and $\mathbb{I}(\psi)$. But if $\mathbb{I}(\phi) = 1$, then $\mathbb{I}(\neg\phi) = 0$, and given that, by assumption, $\mathbb{I}(\phi \to \psi) = 1$, then this must mean that $\mathbb{I}(\psi) = 1$. Hence

[2]It would also make sense to change our restriction that for each $A \in D_1$, there is an $f \in D_2$ such that for each $a \in A, f(a) > 0$ (or, in the formalism of the previous section, that for each $A \in D_2$, $A^+ \cup A^- = D_1$) to $f(a) \neq 1/2$ ($A^+ \cap A^- = \emptyset$). This is because $K3$ is a paracomplete, rather than paraconsistent logic, i.e. it has truth-gaps as opposed to truth-gluts.

modus ponens holds in ST. For proof of the general classicality of ST, see [6]. Given this situation with ST, it might be thought that it would provide no help for us regarding our identity problems, but it turns out this is not the case. Before we come to see why, it will first be useful to address a couple of our target features that we think ST gets right, namely (SI) and (TI).

Recall that in LP, the biconditional between two sentences was strictly true iff they were both strictly true, or both strictly false. This is likewise the case for ST: $\mathbb{I}(\phi \leftrightarrow \psi) = 1$ iff $\mathbb{I}(\phi) = \mathbb{I}(\psi) = 1$ or $\mathbb{I}(\phi) = \mathbb{I}(\psi) = 0$. As we saw with the conditional, we haven't changed any of the inner workings of the logic, only the consequence relation. For identity, then, this will mean that for each predicate P, $\mathbb{I}(Pa) = \mathbb{I}(Pb) = 1$ or $\mathbb{I}(Pa) = \mathbb{I}(Pb) = 0$. If we assume a model where $\mathbb{I}(Pa) = 1$, then, as we must do for (SI), it will follow that $\mathbb{I}(Pa) = \mathbb{I}(Pb) = 1$. Hence (SI) holds for ST. Likewise, with (TI), if we assume $a = b$ and $b = c$ are both value 1, then Pa and Pb and Pc must all have value 1 or all have value 0, individuated according to each predicate. It is clear then that $a = c$ will have value 1, so (TI) holds for ST.

Now let us return to the sorites paradox, and specifically, to the case of the coloured sheets of paper hanging on the wall. One way of representing the argument was as $Ra_{255}, \forall n(a_n = a_{n-1}) \models Ra_0$. This was shorthand for the following elliptical argument:

$$\begin{array}{l} Ra_{255} \\ a_{255} = a_{254} \\ a_{254} = a_{253} \\ \dots \\ a_2 = a_1 \\ a_1 = a_0 \\ \hline Ra_0. \end{array}$$

Recall that in the introduction one way of explaining the standard account's failure at properly dealing with the argument was through the transitivity of identity and the substitutivity of identicals. We have already shown that both these principles hold in ST, that is, (SI) and (TI) are both the case. The suspicion again, then, might be that ST would be ill-equipped to deal with the paradox, much like the standard account. But this is not so. For observe now what role another sort of transitivity plays in ST, namely the transitivity of inference, (ToI).

From (SI) holding in ST, we know that $Ra_{255}, a_{255} = a_{254} \models_{ST} Ra_{254}$, that is we can infer from (some of) our premises that the sheet of paper with color value $\langle 254, 0, 0\rangle$ is red. But note that because ST differentiates sentence values according to their positions, the value of Ra_{254} needn't be the value of Ra_{255}. In short, all the argument implies is that Ra_{254} is tolerantly true. The result of this is that we can't plug Ra_{254} back into another (SI)-type argument to continue the sorites progression, i.e. although $Ra_{254}, a_{254} = a_{253} \models_{ST} Ra_{253}$ holds, given that we can't be sure that $\mathbb{I}(Ra_{254}) = 1$, our first premise is not good enough - it is not strictly true. Contrary to the standard account, then, we can make no judgement regarding the redness of a_{253} in ST. In this way, the sorites argument is cut off very quickly, well before we can reach the troublesome conclusion Ra_0.

What we have said with respect to (SI) also holds for (TI). Although by assumption we know that every adjacent sheet of paper is indistinguishable, and this, by (TI) will mean that sheets adjacent plus one will be identical, we cannot chain the inferences together to render the entire series of sheets identical. For example, although $a_{255} = a_{254}, a_{254} = a_{253} \models_{st} a_{255} = a_{253}$, this only says something relatively weak about $a_{255} = a_{253}$, namely that it is tolerantly true. The result is that we can't go on to make the argument that $a_{255} = a_{253}, a_{253} = a_{252} \models_{ST} a_{255} = a_{252}$, despite this arguments' correctness in ST. Again, the sorites is ended swiftly before we can claim anything like $a_{255} = a_0$, so we avoid the problem associated with the standard account.

In light of these two results, it could be said that ST treats the sorites in as plausible a way as LP. Furthermore, it does so for similar reasons, namely in its use of tolerant truth. We stopped both the never-ending substitutivity of identicals and the never-ending transitivity of identity by relying on this lesser version of truth. More precisely, ST relies on tolerant truth to make it so that (ToI) fails in ST: for some formulas ϕ, ψ, χ, we can have it so that $\phi \models_{ST} \psi$, and $\psi \models_{ST} \chi$, without having $\phi \models_{ST} \chi$.[3] Another way in which we can distinguish LP from ST is that LP cuts off sorites-like progressions one step sooner than ST does. This would mean that ST wouldn't be able to deal

[3]Strictly speaking, (ToI) does not directly address the sheets of paper example as we have represented it, as (ToI) only concerns singletons while our representation implicitly assumed sets with multiple formulas. A more accurate way of putting the substitutivity example (transitivity of identity would be similar) would be that: $Ra_{255} \wedge a_{255} = a_{254} \wedge a_{254} = a_{253} \models_{ST} Ra_{254} \wedge a_{255} = a_{254} \wedge a_{255} = a_{253}$ and $Ra_{254} \wedge a_{255} = a_{254} \wedge a_{255} = a_{253} \models_{ST} Ra_{253}$, but $Ra_{255} \wedge a_{255} = a_{254} \wedge a_{254} = a_{253} \not\models_{ST} Ra_{253}$ (with $=$ taking precedence over the associative $\wedge$). This is clearly a violation of (ToI), and it is also implicit in the reasoning we gave in the example, but it also a little cumbersome to state and hence obscure.

well with a simple three object situation, a situation which LP can handle. This is only a minor inconvenience: usually, sorites situations require many objects. We said before that with the red sheets example, it would make empirical sense to allow (SI) and (TI) a number of times before we held them to fail. Just as with LP, ST can do this: we can have many objects all equal each other in the strong sense required for ST, only for identity to break down somewhere in the middle.

Despite ST's favourable qualities (its acceptance of (SI) and (TI), and its plausible treatment of the sorites) it does suffer from some serious drawbacks, namely in the way that it deals with our vague target features, i.e. (NCO) and (NVI). The premises of these arguments can never be strictly true in the version of ST we have presented. It is impossible, for example, that both $\mathbb{I}(a = b) = 1$ and $\mathbb{I}(a \neq b) = 1$ for any MV_2-model. It will follow that problems like Prior's amoeba example will not find a resolution in the version of ST we have presented. will not fit the more standard definition that $\phi, \psi \models \bot$. We did see however that ST relies on tolerance, and hence, plausibly, paraconsistency, in order to deal with the sorites then. For this reason, we might call ST a para-paraconsistent logic. Although ST is not a paraconsistent logic in the sense that inconsistencies lead to everything in ST, it does allow for inconsistent models in principle. It should also be stressed that the version of ST presented here was only a limited one. In [10] a second version of identity ($\approx$) is also defined, one that is better suited to matters of vagueness. This version of identity *is* capable of allowing models for contradictory premises. Furthermore, it also allows for (SI). Unfortunately, $\approx$ is not transitive, and so, in this respect is more similar to the identity of LP. (Given the non-transitivity of inference of ST though, it can be used to produce some different results.) Because we were looking for versions of identity for which transitivity would hold, we need not look at this version of identity, nor indeed the strict-tolerant approach any further. The main lesson to take from this chapter is that we can keep hold of some consistency, while still being able to handle identity problems. We will try and use this as inspiration to form our third and final logic of the next chapter.

Chapter 4

A Pragmatic Approach

4.1 Towards a Pragmatic Approach

Up to this point, we've seen that we can use paraconsistent logics extrapolated to a second-order to give an account of identity that is better suited to deal with identity problems in the context of vagueness than first- or second-order logic. We also saw, however, that neither *LP* nor *st* were fully capable of capturing our natural intuitions regarding identity, and that specifically, in neither would all our five target features hold true. The aim of this chapter will be to present a logic *prpr* (pronounced "prag - prag") that *will* provide such an account of identity. Before we come to *prpr* in detail, it might first be useful to dwell upon some thoughts that we haven't yet touched upon in detail, namely default assumptions.

One such default assumption that is often spoken about is that of consistency: where possible, we should try our best to give consistent readings. This is no doubt an appealing idea. After all, we were not compelled towards non-classical logics in the first place because classical logic was particularly bad at dealing with consistent scenarios, but rather because it could not handle inconsistent ones. It might make sense then that where everything behaves well, and there is no vagueness, we should fall back on the assumption of consistency. With *LP* we saw evidence of this approach in at least two places: once concerning the non-transitivity of identity, and also the failure (and simultaneous success) of self-identity. In both cases, the cognitive shock brought about by the failure of these familiar laws was tempered by the idea that whenever our models are consistent, identity is indeed transitive and reflexive; it is only when we are dealing with vague objects that anything out of the ordinary occurs. We think that there is an element of truth in the default assumption of consistency, namely that when it is explicitly stated that some such thing is consistent (e.g. "John is tall"), we should not be able to infer vagueness (for example that John is both tall and not tall). The same can be said for the inference from "John is not tall" to "John is vaguely tall".[1] In the absence of information indicating consistency however, we feel that the impartial approach of allowing for inconsistent readings as much as we do consistent ones makes the most sense: if we haven't specifically mentioned whether an object is vague or not, we shouldn't by default rule either case out. We think this, in short, because of the apparent prevalence of inconsistency in the world. If we were to fully embrace the default assumption of consistency and reject inconsistent readings, we would be denying this apparent truth.

We can take our example non-inferences of consistency to vagueness in relation to another default assumption: pragmatism. Suppose you are presented with a disjunction, say, "The world is beautiful or there is no God". Of course, classically, such a disjunction needn't preclude the corresponding conjunction. Just because it has been said that it is true that the world is beautiful or it is true that there is no God doesn't meant that it can't be the case that the world is beautiful and there is no God. From a certain perspective though, such a reading doesn't seem quite right. For wouldn't it be the case, that if the speaker really wanted to allow for the conjunction they would have made that clear? Shouldn't we assume by default the weaker statement (i.e. the disjunction) if the speaker has chosen not to say the stronger one (the conjunction)? If we answer "yes", we are said to allow for scalar implicatures, which are often said to follow from standard Gricean principles[2] We will use this idea, but within the context of vagueness. Earlier, we saw that *prpr* should not allow inferences from, say, "John is tall" to "John is vaguely tall" because if we are given indication of consistency, this should be kept to, by default, and inconsistent readings should be ruled out. From a pragmatic perspective we can reach the same conclusion. For if a speaker informs you that John is tall, then they have, under the default pragmatic assumption, implicitly rejected telling you that they are both tall and not tall. The very fact that they didn't say the stronger statement thus indicates that John must simply be tall, or better put, that "John is tall" is strictly true. Conversely, by analogous reasoning, we think that if an assertion is made to the effect that John is vaguely tall, it would not be correct

[1] There is some empirical evidence for what we are saying here, for example, [40], [1], [9], [36].

[2] See, for example, [15].

to infer that he is, after all, strictly tall. That is, we should have it that $Tj \wedge \neg Tj \not\models_{prpr} Tj$. Allowing for situations such as this will mean that *prpr* will be quite unlike any of the logics we have seen so far - specifically, it will be non-monotonic, i.e. $\Gamma \models_{prpr} \{\phi\} \not\Rightarrow \Gamma \cup \psi \models_{prpr} \phi$. Unsurprisingly, this will have big consequences for how identity behaves in *prpr*. We will see how we can use it to our advantage in dealing with identity problems, for example giving us a novel way of dealing with the sorites and Theseus' Ship paradoxes.

To explicate *prpr*, we will not use one of the formalisms already introduced. Instead, we will take a more "fine-grained" approach to truth, making use of a more general formalism, originally found in van Fraassen's [42]. That this formalism is more general will become clear as we go along, as will the notion of fine-grainedness. As a starting point, we must return to a central question for logic, namely, "Under what circumstances should we consider a sentence to be true?", i.e. what are a sentence's truth-conditions? For van Fraassen, the answer is that sentences are made true by facts. To put it more accurately, sentences have truth-makers, which are sets of facts. Facts are themselves sets, built up from the most primitive elements, namely elements of the *state of affairs*, which intuitively we can think of as the collection of possible situations. The set of possible worlds will be the set of maximal subsets of the state of affairs, that is, the set of all subsets which include at least $\mathbf{p}$. For each atomic sentence p, there is a corresponding element in the state of affairs $\mathbf{p}$. A truth-maker for p is then defined as all the sets that contain the state of affairs $\mathbf{p}$. For complex sentences, let's first look at negated formulas. The first thing we will need is false-makers, that is, facts that make sentences false. We can then define the truth-makers for a negated formula as simply the set of false-makers for the corresponding non-negated formula. Where p is atomic, we can think of at least two different ways of dealing with the truth-makers of $\neg p$, i.e. the false-makers of p. The first would be to say that there is some (positive) state q in the state of affairs that rules out p. The second would be to say that there are negative states in the state of affairs. Van Frassen chooses the latter of these two options. There is thus an intimate connection between the set of literals in a language and the state of affairs: for each atomic sentence p, there is a member of the state of affairs that makes it true: $\mathbf{p}$, and a member that makes it false: $\overline{\mathbf{p}}$. For the truth-makers for conjunctions, naturally enough, we consider conjunctions of facts. To say for example that a conjunction of atomic sentences p and q is true, is to say that its truth-makers must consist of a (conjunctive) fact containing both $\mathbf{p}$ and $\mathbf{q}$, that is, $\{\{\mathbf{p}, \mathbf{q}\}\}$ is its set of truth-makers. For disjunctive sentences, intuitively it makes sense that they may have more than one truth-maker. For example, $p \vee q$ can be made true either by the fact

$\{\mathbf{p}\}$ or $\{\mathbf{q}\}$, and hence its truth-makers come out as $\{\{\mathbf{p}\}, \{\mathbf{q}\}\}$. To generate the false-makers for conjuctions and disjunctions, we can make use of the familiar de Morgan rules to say that a conjunction's false-makers will be a disjunction of the conjuncts' false-makers, and a disjunction's false-makers will be the conjunction of the disjuncts' false-makers. We can extend this to account for higher-order sentences in a natural way, namely that universal statements will be considered as (possibly infinite) conjunctions, and existential statements as disjunctions.

4.2 Formalising *prpr*

We will now spell out what was described informally above.[3] We begin with a definition for a state of affairs, and what our possible worlds are.

Definition of the state of affairs (SOA):

For each atomic sentence (p) of the language $: \mathbf{p}, \overline{\mathbf{p}} \in SOA$ and $\overline{\overline{\mathbf{p}}} = \mathbf{p}$.

Definition of the set of possible worlds (W):

$$W = \{w \in \mathcal{P}(SOA) \mid \forall \mathbf{p} \in SOA : \mathbf{p} \in w \text{ or } \overline{\mathbf{p}} \in w\}.$$

We can now state the simultaneous recursive definition for truth- and false-makers for a formula ϕ, denoted $T(\phi)$ and $F(\phi)$ respectively. As is perhaps already clear, this framework closely resembles disjunctive normal form. Note, by $\otimes$ we will have in mind a set operation similar to the Cartesian product, but rather than generating a set of ordered pairs of two sets, $\otimes$ will generate an (unordered) set, i.e. $A \otimes B = \{X \cup Y \mid X \in A \text{ and } Y \in B\}$.

Definition for propositional truth- and false-makers:

$$\begin{aligned} T(p) &= \{\{\mathbf{p}\}\} & F(p) &= \{\{\overline{\mathbf{p}}\}\} \text{ (where } p \text{ is atomic.)} \\ T(\neg\phi) &= F(\phi) & F(\neg\phi) &= T(\phi) \\ T(\phi \wedge \psi) &= T(\phi) \otimes T(\psi) & F(\phi \wedge \psi) &= F(\phi) \cup F(\psi) \\ T(\phi \vee \psi) &= T(\phi) \cup T(\psi) & F(\phi \vee \psi) &= F(\phi) \otimes F(\psi) \end{aligned}$$

[3] The following account shares many common features to those of [11] and [12]. One notable distinction is in our focus on (general) identity as opposed to the more specific identity-with-respect-to-predicate relation, i.e. the similarity relation $\sim_P$.

For example then, $T(p\wedge(q\vee r)) = T(p)\otimes(T(q\vee r)) = T(p)\otimes(T(q)\cup T(r)) = \{\{\mathbf{p},\mathbf{q}\},\{\mathbf{p},\mathbf{r}\}\}$, and $T(p\vee\neg p) = T(p)\cup F(p) = \{\{\mathbf{p}\},\{\overline{\mathbf{p}}\}\}$. (Note that, in this paper, false-makers will only be useful in so far as they are needed to construct our truth-makers, so I will avoid giving any direct examples of them.)

As before, $\phi \rightarrow \psi$ will be defined standardly as $\neg\phi \vee \psi$, so for example, $T(p \rightarrow q) = \{\{\overline{\mathbf{p}}\},\{\mathbf{q}\}\}$. Likewise, the bi-conditional will be considered as the conjunction of the conditional and its converse, that is, $\phi \leftrightarrow \psi$ iff $(\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi)$. In this latter case, it follows that for some atomic sentences p and q, $T(p \leftrightarrow q) = \{\{\mathbf{p},\mathbf{q}\},\{\overline{\mathbf{p}},\overline{\mathbf{q}}\},\{\mathbf{p},\overline{\mathbf{p}},\},\{\mathbf{q},\overline{\mathbf{q}}\}\}$. This will be of particular interest when we come to defining identity later in this chapter. For the moment, it will suffice to point out that the truth-makers for the bi-conditional connecting atomic sentences include facts whereby each atomic sentence is vague (- for the example given, the sets $\{\mathbf{p},\overline{\mathbf{p}}\}$ and $\{\mathbf{q},\overline{\mathbf{q}}\}$).

For the quantifiers, we will need a way to interpret (atomic) sentences with predicate constants and individual constants. We will assume for simplicity that each element $d \in D_1$ and each $A \in D_2$ has a unique name in the language: $\underline{d}$ and $\underline{A}$ respectively. Where we wish to refer to an element of either domain, we will say it is an element of D. By α we will refer ambiguously to either an individual or predicate variable in the language. As before, we will just be considering monadic predicates, and we will likewise ignore function symbols.

Quantified truth- and false-makers:

$$\begin{aligned} T(\underline{A}\underline{d}) &= \{\{\mathbf{Ad}\}\} & F(\underline{A}\underline{d}) &= \{\{\overline{\mathbf{Ad}}\}\} \\ T(\exists\alpha\phi) &= \bigcup_{d\in D} T(\phi[\alpha/\underline{d}]) & F(\exists\alpha\phi) &= \bigotimes_{d\in D} F(\phi[\alpha/\underline{d}]) \\ T(\forall\alpha\phi) &= \bigotimes_{d\in D} T(\phi[\alpha/\underline{d}]) & F(\forall\alpha\phi) &= \bigcup_{d\in D} F(\phi[\alpha/\underline{d}]) \end{aligned}$$

So for example then, if $D_1 = \{a, b\}$, $T(\exists x(Px\wedge\neg Px)) = \{\{\mathbf{Pa},\overline{\mathbf{Pa}}\},\{\mathbf{Pb},\overline{\mathbf{Pb}}\}\}$, and if $D_1 = \{a, b\}$ and $D_2 = \{P, Q\}$ then $T(\forall X(Xa\vee Xb)) = \{\{\mathbf{Pa},\mathbf{Qa}\},\{\mathbf{Pb},\mathbf{Qb}\}\}$.

Truth- and false-makers can be thought of as providing a fine-grained semantic interpretation, which is really just another way of saying that they can allow us to define notions that are more nuanced than we would normally be able to. One such notion will be be of critical importance to *prpr*. We can use truth- and false-makers together with our worlds to define a wide variety of different logics. For example, to recover the standard truth-conditional semantics for a sentence ϕ, $[\![\phi]\!]$, we limit our worlds to be consistent (as well as maximal), and then define $[\![\phi]\!]$ to be the set of worlds which have a truth maker, i.e. $[\![\phi]\!] = \{w \in W \mid \exists f \in T(\phi) : f \subseteq w\}$. Standard logical consequence would hold between a set of premises Γ and a conclusion ϕ just in case $\bigcap_{\gamma\in\Gamma}[\![\gamma]\!] \subseteq [\![\phi]\!]$.

Where our worlds are simply maximal, and not necessarily consistent (i.e. as we defined W at the beginning of this section), the set of worlds for which ϕ has a truth-maker becomes the set of tolerant worlds for ϕ. That is, $[\![\phi]\!]_t = \{w \in W \mid \exists f \in T(\phi) : f \subseteq w\}$. We can then define strict truth in terms of this: $[\![\phi]\!]_s = [\![\phi]\!]_t \cap MC$, where MC is the set of maximally consistent sets of W. Put another way, $[\![\phi]\!]_s = \{w \in W \mid \exists f \in T(\phi) : f \subseteq w : \mathbf{p} \in w \text{ and } \overline{\mathbf{p}} \notin w\}$. (It should be noted that $\mathbf{p}$ here, as elsewhere, means any literal, not necessarily a positive state of affairs. Otherwise, this wouldn't be right.) From this we can define the logics seen in the previous chapter in the following way:

$$\Gamma \models_{LP} \phi \text{ iff } \bigcap_{\gamma\in\Gamma}[\![\gamma]\!]_t \subseteq [\![\phi]\!]_t,$$

$$\Gamma \models_{st} \phi \text{ iff } \bigcap_{\gamma\in\Gamma}[\![\gamma]\!]_s \subseteq [\![\phi]\!]_t.$$

Broadly speaking, *prpr* will be constructed in the same way: a pragmatic interpretation of ϕ will restrict our worlds to those which contain truth-makers for ϕ, given some further conditions; and *prpr* consequence will be said to hold between a set of premises Γ and a conclusion ϕ iff the intersection of the pragmatic Γ worlds are all included in the pragmatic ϕ ones. As to what counts as "further conditions", we will make use of the notion of minimally inconsistent worlds.

4.3 Pragmatic Interpretation

The task of making precise the notion of minimal inconsistency essentially boils down to the problem of creating an ordering amongst worlds with respect to their consistency. In [**?**] (pp. 7-8), this was carried out by defining a world w as less inconsistent than another v iff v has every inconsistency of w and more, i.e. that $w < v$ iff $\{\mathbf{p} \in SOA \mid \mathbf{p} \in w \text{ and } \overline{\mathbf{p}} \in w\} \subset \{\mathbf{p} \in SOA \mid \mathbf{p} \in v \text{ and } \overline{\mathbf{p}} \in v\}$. A minimally inconsistent world w_1 then, is simply one for which there is no other w_2 such that $w_2 < w_1$. The truth-conditions for a sentence ϕ can then be defined as the set of minimally inconsistent worlds for which ϕ has a truth-maker in w, i.e. $\{w \in W \mid \exists f \in T(\phi) : f \subseteq w \text{ and } \neg\exists v \in W : f \subseteq v \text{ and } w < v\}$. Although not without its merits, this approach has the unfortunate consequence of unfairly discounting what we consider to be harmlessly inconsistent worlds. A notable example discussed in [**?**] concerns the sentence "John is vaguely tall or Mary is rich". It is not hard to see why such a sentence might be problematic: we have a disjunction of an inconsistency (John being vaguely tall) and a strict truth (Mary being tall), and minimal inconsistency, as defined above, will seek to minimise inconsistency, avoiding cases whereby the first disjunct is true. In slightly more detail, suppose we have two worlds, $w_1 = \{Tj, \neg Tj, \overline{Rm}\}$, and $w_2 = \{Tj, Rm\}$. Both worlds contain at least one truth-maker for the whole sentence ($\{Tj, \overline{Tj}\}$ and $\{Rm\}$, respectively), and so are initially included by the interpretation, but because $w_2 < w_1$ (because $\emptyset \subset \{Tj, \overline{Tj}\}$), a world such as w_1 will never emerge from the above interpretation. In fact, the interpretation above will render the whole sentence as being equivalent to Mary being tall.[4] We think this wrong, intuitively, simply because there seem to be possible worlds in which Mary is in fact not tall, and yet the whole sentence still true pragmatically, namely when the other disjunct is true, and John is vaguely tall. For this reason, we will reject Priest's definition of less inconsistent worlds in favour of the following definition.

<u>Definition of a world (w) being less inconsistent than another (v):</u>

$$v <_f w \text{ iff } \{p \in SOA \mid p \in f \text{ and } \overline{p} \in v\} \subset \{p \in SOA \mid p \in f \text{ and } \overline{p} \in w\}.$$

We can see that inconsistency is thus relativised according to a specific truth-maker f. As such, we will not face the problem we did above concerning the "John is vaguely tall or Mary is rich" problem. If we consider again our two worlds w_1 and w_2, each has a different truth-maker, and because our definition of being less inconsistent is relativised to them separately, they will each be minimally inconsistent in their own right. This specific example will be treated in more detail once we have concluded formalising *prpr*.

<u>Definition of *Pragmatic Interpretation*:</u>

$$PRAG(\phi) = \{w \in W \mid \exists f \in T(\phi) : f \subseteq w \text{ and } \neg\exists v \in W : f \subseteq v \text{ and } v <_f w\}.$$

To re-iterate, $PRAG(\phi)$ looks for worlds for which ϕ has a truth-maker f which are as consistent as possible with respect to f. If for a given world w, there is another v that also makes ϕ true, and is less inconsistent than

[4] See [**?**] for more details on this.

w, then w will be excluded and not emerge from our pragmatic interpretation.

Where we are only considering single-premised sequents, concluding the logic would be simply enough. We could let $\phi \models_{prpr} \psi$ hold iff $PRAG(\phi) \subseteq PRAG(\psi)$. Where we have more than one premise however, the situation is more tricky. This is because if we define things in the standard way, i.e. that $\Gamma \models_{prpr} \phi$ iff $\bigcap_{\gamma\in\Gamma} PRAG(\gamma) \subseteq PRAG(\phi)$, we will find that we will not adequately be able to deal with relatively inconsistent premises. Consider, for example, where $\Gamma = \{Pa, \neg Pa\}$. As we have defined things, $PRAG(Pa)$ and $PRAG(\neg(Pa))$ will be disjoint, and so $PRAG(\Gamma)$ will be empty, resulting in an incorrect reading[5]. Instead we would like the pragmatic interpretation for premises to behave much in the same way as $PRAG$ deals with conjunctions, so that in the example given, $PRAG(\Gamma)$ behaved in much the same way as $PRAG(Pa \wedge \neg Pa)$. The following definition allows for this.

Definition for multi-premised $PRAG$:

$$PRAG(\Gamma) = \begin{cases} \bigcap\limits_{\gamma\in\Gamma} PRAG(\gamma, W), & \text{if this is non-empty;} \\ \bigcap\limits_{\gamma\in\Gamma} PRAG(\gamma, [\![\Gamma]\!]_t), & \text{otherwise,} \end{cases}$$

where,

$$PRAG(\phi, W') = \{w' \in W' \mid \exists f \in T(\phi) : f \subseteq w' \text{ and } \neg\exists v \in W' : f \subseteq v \text{ and } v <_f w\},$$

and,

$$[\![\Gamma]\!]_t = \bigcap_{\gamma\in\Gamma} [\![\gamma]\!]_t.$$

To see why, consider again the example where $\Gamma = \{Pa, \neg Pa\}$. As previously stated, if we perform $PRAG$ for all worlds, $PRAG(\Gamma) = \emptyset$, hence our definition forces us to reinterpret Γ, instead considering the worlds in which every element of Γ is tolerantly true. The result will be that $PRAG(\Gamma) = PRAG(Pa \wedge \neg Pa)$ which is what we wanted. (Note that this is not generally the case for *prpr*. Later in this chapter we will see a counterexample to it related to our target feature (NVI)) We can now state what it means for a sequent to hold in *prpr*.

[5]Note that $PRAG(\emptyset)$ is very different from $PRAG(\Gamma) = \emptyset$. While the latter would allow for any sequent holding, simply because the emptyset is a subset of every set, the former can be understood in the standard way regarding logical consequence, as providing no demands on models for the premises, hence forcing the pragmatic interpretation of the conclusion needing to be equal to W.

Definition of *prpr* consequence:

$$\Gamma \models^{prpr} \phi \text{ iff } PRAG(\Gamma) \subseteq PRAG(\phi).$$

This concludes the basics of *prpr*. Before coming to how *prpr* deals with identity, it might first be useful to see how it deals with sequents involving simpler formulas, including some that we considered informally at the beginning of this chapter.

One of the first ideas we considered, was that where consistency has explicitly been mentioned, we should not then allow for inconsistency. For example, it should hold that $Pa \not\models_{prpr} Pa \wedge \neg Pa$. While this is the case for many logics - including LP - simply because there exists a model whereby the premise is true and the conclusion false, the situation is more extreme in *prpr*: there is simply no world $w \in PRAG(Pa)$ such that $w \in PRAG(Pa \wedge \neg Pa)$. It is in this sense that inconsistency is said to be ruled out, given evidence of consistency. With the tolerant approach (recall that LP can be thought a tolerant-to-tolerant logic), we could not reach such an extreme position because the tolerant worlds for Pa would be any set containing Pa (i.e. $T(Pa) = \{\mathbf{Pa}\} \subseteq$ those sets), and so worlds w for which $\{\mathbf{Pa}, \overline{\mathbf{Pa}}\} \subseteq w$ could not necessarily be ruled out. Put in more intuitive terms, the premise Pa, when understood tolerantly, may only be half true, that is $Pa \wedge \neg Pa$ might hold, in which case $Pa \wedge \neg Pa$ would be true. But with $PRAG$, while $\{\mathbf{Pa}\} \subseteq w$ for any world w containing $\mathbf{Pa}$, any world that also contains $\overline{\mathbf{Pa}}$ will be more inconsistent than w (assuming there it has no other truth-maker), and so be excluded by our definition. It follows that the $PRAG$ worlds for Pa will only be the consistent ones - at least in terms of the truth-maker $\mathbf{Pa}$. Or put another way, the pragmatic interpretation of Pa says that Pa must be strictly true. As for other, non-explicitly mentioned atomic sentences, $PRAG$ does not force us to make any judgments as to their consistency because of the way we relativised inconsistency according to truth-makers. It will thus be the case that for any $q \neq Pa$, the pragmatic worlds will include those in which q are strictly true, strictly false, or both true and false, equally. It was this that we were giving indication of in the beginning of the chapter, saying that in general, *prpr* would hold no default assumption of consistency.

We also mentioned the perhaps more surprising result that $Pa \wedge \neg Pa \not\models_{prpr} Pa$, an example that shows the non-monotonic nature of *prpr*. In the light of what has been said, this is perhaps less surprising. All of the worlds in $PRAG(Pa \wedge \neg Pa)$ will be supersets of $\{\mathbf{Pa}, \overline{\mathbf{Pa}}\}$, while none of the same supersets will be found in $PRAG(Pa)$ because it limits Pa to being strictly true. This gives evidence for the fact that statements concerning inconsistency (e.g. $Pa \wedge \neg Pa$) behave in *prpr* very differently from how they do in LP and ST, and that specifically we are better able to map our natural intuitions (such as the idea that we can't conclude a consistency from an inconsistency) because of this.

As a final example, we will consider in more detail the "John is vaguely tall or Mary is rich" problem, with the aim of showing that we cannot infer from this sentence that Mary is rich. To this end, it will suffice to show that there is a model for which the premise $(Tj \wedge \neg Tj) \vee Rm$ is pragmatically true, while the conclusion Rm is pragmatically false, or more accurately, that there is a world in $PRAG((Tj \wedge \neg Tj) \vee Rm)$ that is not in $PRAG(Rm)$. In taking a more exhaustive approach, we will perhaps get a clearer sense of how the elements of *prpr* we have thus discussed so far come together.

Let $SOA = \{\mathbf{Tj}, \overline{\mathbf{Tj}}, \mathbf{Rm}, \overline{\mathbf{Rm}}\}$. It follows that we will have 9 possible worlds, the set of maximal subsets of SOA, which we will label w_1 through w_9. Let $\phi = (Tj \wedge \neg Tj) \vee Rm$. Recall that $T(\phi) = \{\{\mathbf{Tj}, \overline{\mathbf{Tj}}\}, \{\mathbf{Rm}\}\}$. The following table shows which worlds satisfy the two conditions of $PRAG(\phi)$.

world label	world (w)	$\exists f \in T(\phi) : f \subseteq w$?	$\neg\exists v \in w : f \subseteq v$ and $w <_f v$?
w_1	$\{Tj, Rm\}$	yes: $\{Rm\}$	yes
w_2	$\{Tj, \overline{Rm}\}$	no	-
w_3	$\{Tj, Rm, \overline{Rm}\}$	yes: $\{Rm\}$	no: $w_3 <_{\{Rm\}} w_1$
w_4	$\{\overline{Tj}, Rm\}$	yes: $\{Rm\}$	yes
w_5	$\{\overline{Tj}, \overline{Rm}\}$	no	-
w_6	$\{\overline{Tj}, Rm, \overline{Rm}\}$	yes: $\{Rm\}$	no: $w_6 <_{\{Rm\}} w_1$
w_7	$\{Tj, \overline{Tj}, Rm\}$	yes: $\{Tj, \overline{Tj}\}, \{Rm\}$	yes: either
w_8	$\{Tj, \overline{Tj}, \overline{Rm}\}$	yes: $\{Tj, \overline{Tj}\}$	yes
w_9	$\{Tj, \overline{Tj}, Rm, \overline{Rm}\}$	yes: $\{Tj, \overline{Tj}\}, \{Rm\}$	yes: $\{Tj, \overline{Tj}\}$

We can see that some worlds, namely w_2 and w_5 fail to have any truth-maker for ϕ. These are excluded immediately. There are other worlds that do contain at least one ϕ truth-maker, but that fail to be minimally inconsistent worlds. For example, w_3 is a superset of the truth-maker $\{\mathbf{Rm}\}$, but w_1 is more consistent than w_3 in terms of it (because $\emptyset \subset \{\mathbf{Rm}, \overline{\mathbf{Rm}}\}$), it is thus excluded from $PRAG(\phi)$. Similarly, w_6 is excluded. With every other world, our two $PRAG$ conditions are met. Of particular note are w_7 and w_9. w_7 is minimally inconsistent in terms of either of the truth-makers, while in the latter, only $\{Tj, \overline{Tj}\}$ (because $w_1 <_{\{\mathbf{Rm}\}} w_9$). It follows that

$PRAG(\phi) = \{w_1, w_4, w_7, w_8, w_9\}$.

The table also makes it fairly clear which worlds emerge from $PRAG(Rm)$. They will be the worlds which have (just) Rm, i.e. the worlds that don't also have $\overline{\mathbf{Rm}}$, i.e. $PRAG(Rm) = \{w_1, w_4, w_7\}$. Clearly then, there is at least one world, for example w_8 which is a $PRAG$ world for ϕ and not for Rm. It follows that $PRAG(\phi) \nsubseteq PRAG(Rm)$ and $(Tj \wedge \neg Tj) \vee Rm \not\models_{prpr} Rm$, which is what we wanted to show.

4.4 A Return to Identity

In the two previous chapters we saw that identity could be defined in the familiar Leibnizian way, i.e. $a = b$ iff $\forall X(Xa \leftrightarrow Xb)$. This was somewhat appealing in that it could be argued that any non-standard behaviour of identity was a result of the underlying (non-classical) logics from which it emerged, and not *per se* from identity itself being defined non-standardly. As mentioned previously though, the truth-makers for the bi-conditional (say, $p \leftrightarrow q$) include cases of vagueness (e.g. the sets $\{p, \overline{p}\}$ and $\{q, \overline{q}\}$). Were we to define identity in the Leibnizian way then, it would follow that $T(a = b) = T(\forall X(Xa \leftrightarrow Xb)) = \bigotimes_{A \in D_2}\{\{\mathbf{Aa}, \mathbf{Ab}\}, \{\overline{\mathbf{Aa}}, \overline{\mathbf{Ab}}\}, \{\mathbf{Aa}, \overline{\mathbf{Aa}}\}, \{\mathbf{Ab}, \overline{\mathbf{Ab}}\}\}$. The presence of these two final sets would generally be problematic in light of our $PRAG$ definition because it would go against our insistence that where there is evidence of consistency, we should not be able to derive inconsistency. By having, for example, $\{\mathbf{Pa}, \overline{\mathbf{Pa}}\}$ as a truth-maker for identity statements, worlds which are subsets of $\{\mathbf{Pa}, \overline{\mathbf{Pa}}\}$ would automatically emerge from our $PRAG$ definition as they would contain a truth-maker and be minimally inconsistent worlds with respect to this truth-maker. However, from the pragmatic perspective, there is every reason to think that asserting an identity statement is an assertion of consistency, much like it is for, say, an assertion of an atomic sentence. In fact, were we to commit ourselves to the above definition of identity, the result would be disastrous in reference to our target features. For example, if we did not implicitly rely on the consistency of an asserted identity statement, we would be quite unable to properly deal with the substitution of identicals. Because of this, we will have to embrace an alternate definition, one that by fiat removes these latter two sets.

<u>Definition for the truth- and false-makers of Identity</u>:

$$T(a = b) = \bigotimes_{A \in D_2} \{\{\mathbf{Aa}, \mathbf{Ab}\}, \{\overline{\mathbf{Aa}}, \overline{\mathbf{Ab}}\}\}$$

$$F(a=b)=\bigcup_{A\in D_2}\{\{\mathbf{Aa},\overline{\mathbf{Ab}}\},\{\overline{\mathbf{Aa}},\mathbf{Ab}\}\}.$$

It is easy to see that, given only consistent worlds, our truth- and false-makers for identity will be just the standard truth-conditional ones. We are, after all, simply saying that for an identity statement to be true we must have it that for every predicate, the conjunction of both predicated objects is true or both are false; and with regards to false-makers, that one predicated object is true while the other false. As Priest puts it, using slightly different terminology, "What is required for Leibniz's Law is that for every predicate, P, Pt_1 and Pt_2 have the same truth value" ([31] p.7). It is in this sense that we think that, as defined, we capture the spirit of the Leibnizian definition, though not, we must concede, the form. (More strictly, it corresponds to the second definition of identity in the introduction as opposed to the first.) This point will further be discussed in the concluding chapter.

4.5 *prpr* and the Target Features

We are now in a position to see how *prpr* deals with our five target features. We will also see what effect this has in dealing with identity problems in the context of vagueness. Firstly, consider (SI), that is, $Pa, a=b \models Pb$.

Let $\Gamma=\{Pa, a=b\}$. Clearly $T(Pa)=\{\{\mathbf{Pa}\}\}$, $T(a=b)=\bigotimes_{A\in D_2}\{\{\mathbf{Aa},\mathbf{Ab}\},$ $\{\overline{\mathbf{Aa}},\overline{\mathbf{Ab}}\}\}$, and $T(Pb)=\{\{\mathbf{Pb}\}\}$. It follows that every world in $PRAG(Pa)$ will contain $\mathbf{Pa}$, but, as we saw before, $\{\mathbf{Pa},\overline{\mathbf{Pa}}\}\not\subseteq w$ for any $w\in PRAG(Pa)$ because if this were the case then these worlds would be more inconsistent than those that contain just $\mathbf{Pa}$. By the same reasoning, $PRAG(Pb)$ will include only worlds in which $\mathbf{Pb}$ is a member, and not also $\overline{\mathbf{Pb}}$. For $PRAG(a=b)$, we know that every world must either be a superset of $\{\mathbf{Pa},\mathbf{Pb}\}$ or $\{\overline{\mathbf{Pa}},\overline{\mathbf{Pb}}\}$. (For example, if $D_1=\{a,b\}$ and $D_2=\{P,Q\}$, $PRAG(a=b)=$ $\{\{\mathbf{Pa},\mathbf{Pb},\mathbf{Qa},\mathbf{Qb}\},$
$\{\mathbf{Pa},\mathbf{Pb},\overline{\mathbf{Qa}},\overline{\mathbf{Qb}}\},\{\overline{\mathbf{Pa}},\overline{\mathbf{Pb}},\mathbf{Qa},\mathbf{Qb}\},\{\overline{\mathbf{Pa}},\overline{\mathbf{Pb}},\overline{\mathbf{Qa}},\overline{\mathbf{Qb}}\}\}$.) From our definition of $PRAG(\Gamma)$, we first interpret Γ according to the set of worlds W, and if the resulting set is non-empty then this is our pragmatic interpretation. Any of those sets in $PRAG(a=b)$ which are supersets of $\{\overline{\mathbf{Pa}},\overline{\mathbf{Pb}}\}$ will be disjoint to those that are supersets of those (just) containing $\mathbf{Pa}$, hence these will be excluded. The only sets that won't be excluded will be those which are supersets of $\{\mathbf{Pa},\mathbf{Pb}\}$. As said previously, $PRAG(Pb)$ will contain supersets of $\{\mathbf{Pb}\}$ (and not also $\{\overline{\mathbf{Pb}}\}$. It follows that $PRAG(\Gamma)\subseteq PRAG(Pb)$, and so $Pa, a=b\models_{prpr} Pb$.

Of course, objects a, b, and predicate P were chosen arbitrarily, and so what was said regarding (SI) also holds in the context of the sorites. That is, where we have some sequence of pair-wise indistinguishable objects, and we assert these identities, together with a claim concerning the $X-$ness of one of these objects, where X is some predicate, we will pragmatically conclude that every object is X too (or symbolically, that $Xa_1, \forall n(a_n = a_{n+1}) \models_{prpr} \forall n(X_n))$. In other words, we find no problem at all with the argument that if a single grain is not a heap, and any addition to a non-heap is also a non-heap, we must pragmatically conclude, given what has been said, that an arbitrarily large number of grains is also not a heap. Recall the specific argument concerning red sheets of paper given in the introduction:

$$\begin{array}{l} Ra_{255} \\ a_{255} = a_{254} \\ a_{254} = a_{253} \\ \ldots \\ a_2 = a_1 \\ a_1 = a_0 \\ \hline Ra_0. \end{array}$$

We likewise no fault with this argument. *prpr* concludes that the final sheet of paper is red, that is, $Ra_{255}, \forall n(a_n = a_{n-1}) \models_{prpr} Ra_0$.

This is not the end of the story though. For with further information of the right sort, the possible worlds made true by $PRAG$ will end up being radically different. Specifically, if at some point it becomes clear that for some object z in the sequence, $\neg Xz$ holds, or in our example - "z *is* a heap" is true, then this will force a radical reinterpretation. Likewise, with the red sheets example, if we come to learn that, say, Ra_0, then we are forced to reinterpret. As to what exact change this introduces, let us consider the so-called "non-absurdity of inconsistent objects" target feature: $Pa, a = b, \neg Pb \not\models_{prpr} \bot$ (NIO). The first thing to say is that where $\Gamma = \{Pa, a = b, \neg Pb\}$, the intersection of the respective $PRAG$ premises will be empty. That is, our premises are relatively inconsistent, the reason being that $PRAG(Pa)$ will generate Pa consistent worlds, $PRAG(a = b)$ will generate $a = b$ consistent worlds, and $PRAG(\neg Pb)$ will generate $\neg Pa$ consistent worlds, and there is simply no world which is a member of all three. We are thus forced to consider the $PRAG$ worlds of the tolerantly Γ worlds, $[\![\Gamma]\!]_t$, as opposed to W. The resulting $PRAG$ worlds will be a set of worlds containing supersets of $\{\mathbf{Pa}, \overline{\mathbf{Pa}}, \mathbf{Pb}\}$, $\{\mathbf{Pa}, \mathbf{Pb}, \overline{\mathbf{Pb}}\}$, or $\mathbf{Pa}, \overline{\mathbf{Pa}}, \mathbf{Pb}, \overline{\mathbf{Pb}}\}$. Or in other words: one (or both) of our objects will be vaguely P. Note that $PRAG(\Gamma)$ will not be empty, that is (NIO) holds in *prpr*.

While (NIO) only concerned two objects, given any number of objects, the situation is the same: one, or possibly any number of objects in a sorites-like sequence must be vague (given the information that two hold inconsistent predicates and every pair-wise object is identical). This means then, that with our red sheets example, given the new information that $\neg Ra_0$ this instantly makes every one of the sheets of paper now borderline red, i.e. $\forall n(Ra_n \wedge \neg Ra_n)$, where $0 \leq n \leq 255$. On the face of it, this might seem to be a rather unsatisfying result. It does, after all, bear some resemblance to the classical problem with the sorites as described at the beginning of this paper, namely that if we have a sequence of pair-wise identical objects, and the first is said to have some predicate, and the last not, then every object in such a sequence must be contradictory. One obvious difference is that inconsistencies in *prpr* do not behave as traditional contradictions. They do not, for example entail anything and everything. Instead they mean something quite specific, that when such and such an object has both a property and its negation, it is vague in terms of that property. The best way of understanding the apparent omni-vagueness that comes out of the *prpr* reading of the sorites then, might be to invoke some notion of meta-vagueness: that sometimes, it might be vague as to what objects are vague or not. This does not seem unreasonable in the context of the sorites, and perhaps likewise for other contexts, including Theseus' Ship. What is perhaps surprising, especially considering on our insistence that from consistency we don't pragmatically infer inconsistency, is that our assertions of strict truth (i.e. that Ra_{255} and $\neg Ra_0$) end up being ignored. Furthermore, no amount of continued assertions of consistency will change the matter. That is, we can insist as much as we care to that, say, object a_{255} really is R, it will not make one jot of difference: it will remain the reading that it is only vaguely red. More will be made of this point in the conclusion that follows this chapter. As a final word, for now, on the sorites, it will be important to mention one such assertion that *will* make a difference, and that is an assertion that one of the objects is vague. For with a further assertion of say, $Ra_{127} \wedge \neg Ra_{127}$, the $PRAG$ worlds that emerge from the premises will then resolve down to only include those which are supersets of the truth-maker $\{\mathbf{Ra}_{127}, \overline{\mathbf{Ra}_{127}}\}$. All other worlds, i.e. the ones where other objects are vague will then fail to come out of the pragmatic interpretation. In other words, once the vagueness has been pinpointed (either one object or more), every other object will then return to being interpreted as strictly true or strictly false. This is a further reason to think that *prpr* deals with the identity sorites (and hence examples like the Ship of Theseus) in a reasonable manner.

We have three target features remaining to take care of: transitive identity (TI), transitivity of inference (TOI), and finally, the non-absurdity of vague identity (NVI). The first follows from a similar argument to the one we saw with (SI). The worlds made true by $PRAG(a = b)$ and the worlds made true by $PRAG(b = c)$ are not disjoint: they will consist of the supersets of $\{\mathbf{Pa}, \mathbf{Pb}, \mathbf{Pc}\}$ or $\{\overline{\mathbf{Pa}}, \overline{\mathbf{Pb}}, \overline{\mathbf{Pc}}\}$. Such worlds will clearly coincide with those of $PRAG(a = c)$, which are supersets of $\{\mathbf{Pa}, \mathbf{Pc}\}$ or $\{\overline{\mathbf{Pa}}, \overline{\mathbf{Pc}}\}$. It follows that $PRAG(\{a = b, b = c\}) \subseteq PRAG(a = c)$, that is, ($TI$) holds. This does not, however, mean that we can't deal with cases where the transitivity of identity is put under strain. In fact, we have already seen hints of this with (NVI). Observe that it is the case that $Pa, a = b, \neg Pb \not\models_{prpr} a = b$, because $PRAG(a = b)$ will only include worlds which are supersets of $\{\mathbf{Pa}, \mathbf{Pb}\}$ or $\{\overline{\mathbf{Pa}}, \overline{\mathbf{Pb}}\}$, while there are no consistent worlds that will be in $PRAG(Pa, a = b, \neg Pb)$. Intuitively, this can be explained: when we utter (just) an identity statement, we are implicitly committing ourselves to consistency (at least in terms of the objects that are said to be identical), while asserting our premises, we are explicitly denying this. Let's look at the truthmakers for an identity statement and its negation: $T(a = b \wedge \neg(a = b)) = \bigotimes_{A \in D_2}\{\{\mathbf{Aa}, \mathbf{Ab}\}, \{\overline{\mathbf{Aa}}, \overline{\mathbf{Ab}}\}\} \bigcup_{B \in D_2}\{\{\mathbf{Ba}, \overline{\mathbf{Bb}}\}, \{\overline{\mathbf{Ba}}, \mathbf{Bb}\}\}$,. Clearly there will always be some worlds that satisfy $PRAG(a = b \wedge \neg(a = b))$, (for example, if $D_1 = \{a, b\}$, and $D_2 = \{P\}$, the world $\{\mathbf{Pa}, \overline{\mathbf{Pa}}, \mathbf{Pb}\} \in PRAG(a = b \wedge \neg(a = b))$). It turns out though that $PRAG(a = b \wedge a \neq b)$ is not quite the same as $PRAG(a = b, a \neq b)$. This is because the former excludes cases whereby both objects are vague. $PRAG$, because of its want to minimise inconsistency, ends up rejecting the case where both objects are vague. It will be the case then that while (NVI) is the case (we can't infer anything from $a = b, a \neq b$), we can't infer that $a = b \wedge a \neq b$. This is therefore a counterexample to the sequent $\phi, \psi \models \phi \wedge \psi$ that we said we'd come to earlier. Although this is perhaps not entirely welcome, there does seem to be a way of making sense of it intuitively. Suppose you have come to learn certain things about one person, and also certain things about a person that you thought was someone else. Perhaps a somewhat concrete example would be that of superman/clark kent. Suppose Lois thought Superman was strong, but that Clark Kent was weak, i.e. not strong. But then, when she found out that Clark Kent and Superman were the same person, to make sense of it, she pragmatically tried to interpret what was going on. The most reasonable assumption, perhaps, is that one of them was lying. So, either Clark wasn't really weak, or Superman wasn't really strong. This at least seems plausible.

Indeed, we might take it to be the most plausible option that Clark Kent is, after all, not strong, and the pragmatic reading we have given allows for this, while LP could not. It should be noted that it is possible to regain an interpretation that does allow both: we can stipulate this as a premise. For example, Lois can say: $Ss, \neg Ss, , \neg Sc, s = c, s \neq c$, and it will follow that both Superman and Clark Kent are vague. Because of this, vague identity (as we have called it) can be made to act exactly as it does in LP. Before, we saw that LP at least plausibly allows us to make sense of Prior's amoeba example. It follows then that *prpr* will be as equally capable in this regard. We can use *prpr* to deal with cases like the amoeba.

Our final target feature, (ToI) follows from the properties of inclusion. Where A, B and C are formulas, clearly if $PRAG(A) \subseteq PRAG(B)$ and $PRAG(B) \subseteq PRAG(C)$, then $PRAG(A) \subseteq PRAG(C)$. Therefore, $A \models_{prpr} B, B \models_{prpr} C \Rightarrow A \models_{prpr} C$. Where we are considering multi-premised sequents, the situation is not complicated any further, because our conclusions are not themselves sets, so $\Gamma \models_{prpr} \phi, \phi \models_{prpr} \psi \Rightarrow \Gamma \models_{prpr} \psi$. Considering multi-conclusion sequents might be source of further work. It would likely be not entirely straightforward, and may also put strain on (TOI) in the more more general case.

Chapter 5

Problems, Further Research, Conclusion

5.1 *prpr* - pros and cons

We can represent the results of the previous chapters (in terms of the target features, at least) in the following table:

	$FOL_=$/ SOL	LP	ST	$prpr$
(SI)	✓	×	✓	✓
(TI)	✓	×	✓	✓
(NIO)	×	✓	×	✓
(NVI)	×	✓	×	✓
(ToI)	✓	✓	×	✓

We have argued that *prpr* thus better suits our intuitions regarding identity, and that, in a sense, it provides us with the best of both worlds with respect to our other two logics: we can deal with vagueness in a way comparible to LP (contra ST), and yet still retain a level of consistency comparible to ST (contra LP). Furthermore, *prpr* provided plausible solutions to our identity problems (the sorites, Theseus' Ship, Prior's amoeba), problems that could not be accounted for by the standard account. We saw that *prpr* was able to do this by channeling a more pragmatic notion of truth, one inspired by how speakers can use language to sometimes imply things without saying them. For example, we saw that *prpr* could make sense of situations where a speaker implicitly avoids inconsistency without explicitly doing so. Such situations were not capturable in LP because it was always possible for sentences involving identity, or atomic sentences to be given inconsistent readings. We also saw that the minimal inconsistency approach taken by *prpr* was more nuanced than it is in other places (Priest's version in somewhere) in that it wouldn't automatically extinguish inconsistency in the presence of consistency (specifically, it provided a meaningful reading of the "Mary is rich or John is vaguely tall" example). We think these are all positive features of *prpr*, and in virtue of them, *prpr* appears to hold some promise. It would be unwise to end our analysis here, though. There are always bound to be *some* percieved problems with any given logic, and *prpr* is of course no different. This will be a suitable place to mention a few problems we might have, with an aim to address them as best we can.

One initial clarification might serve some purpose. This is in regards to how the various target features have been said to hold in *prpr*. The criticism could be put like this. Though technically (SI), say, holds in *prpr*, it only does so in a very limited sense. That is, it is really the case that substitutivity in *prpr* is the same as substitutivity in LP, in that, in both, it fails when we are in inconsistent situations. While this is loosely correct, there is nonetheless a crucial difference between *prpr* and LP, and that is that for *prpr*, it as assumed by default that substitutivity, say, holds, and likewise for transitivity. This is an important difference because it means that in the absence of inconsistent information, *prpr* will get the intuitively right answer. For example, with our version of the sorites paradox, this stresses that there is nothing wrong, in itself, with the argument. If you are told that some object is always identical to the last version of itself, then the intuitive reading is that this object is unchanging - it doesn't change properties. It is only in light of new, inconsistent information (i.e. that one of the sheets isn't red, or that the object now has a property that it didn't have) that we need to go back and change our reasoning. *prpr* clearly allows for this, while LP can't. Another way of describing *prpr*'s treatment of (SI) and (TI) is that (SI) and (TI) are *soft constraints* in *prpr*.

An issue that *prpr* arguably gets wrong is self-identity. We noted in the introduction that together with symmetry and transitivity, reflexivity is often perceived to be one of the defining features of identity, as the smallest equivalence relation. Indeed, however we define identity, its reflexivity appears undeniable. Even Wittgenstein, who is skeptical about a number of issues concerning identity (for example, that it is a relation) appears convinced that everything is self-identical:

> *For identity we seem to have an infallible paradigm: namely, in the identity of a thing with itself. I feel like saying: 'Here at*

> *any rate there can't be different interpretations. If someone sees a thing, he sees identity too.'*[1]

We saw previously that (TI) holds for *prpr*, and the situation with symmetry is much the same: it is clear that $T(a = b) = \bigotimes_{A \in D_2}\{\{Aa, Ab\}, \{\overline{Aa}, \overline{Ab}\}\} = \bigotimes_{A \in D_2}\{\{\mathbf{Ab}, \mathbf{Aa}\}, \{\overline{\mathbf{Ab}}, \overline{\mathbf{Aa}}\} = T(b = a)$, and so $PRAG(a = b) = PRAG(b = a)$, and $a = b \models_{prpr} b = a$. Contrary to our intuitions, though, $\not\models_{prpr} a = a$. Identity is not reflexive in *prpr*. If we look at the truth-makers for $a = a$ where we have only one predicate and one object a for simplicity, it is easy to see why. Where P is our only predicate, $T(a = a) = \{\{\mathbf{Pa}\}, \{\overline{\mathbf{Pa}}\}\}$. It will follow, then, that $PRAG(a = a)$ will not include the world/set $\{\mathbf{Pa}, \overline{\mathbf{Pa}}\}$ because it is more inconsistent than either of the ones in $T(a = a)$ in terms of them. That is, $\{\mathbf{Pa}\} <_{\{\mathbf{Pa}\}} \{\mathbf{Pa}, \overline{\mathbf{Pa}}\}$ and $\{\overline{\mathbf{Pa}}\} <_{\{\overline{\mathbf{Pa}}\}} \{\mathbf{Pa}, \overline{\mathbf{Pa}}\}$. It follows that $PRAG(a = a) \neq W$ (the set of all worlds), and hence $\not\models_{prpr} a = a$. Informally, we can think of the failure as resulting from our commitment to the principle that uttering an identity statement is much like uttering an atomic statement. In both, we assume that the speaker is implicitly making a claim of consistency. Given that not all our worlds are consistent, then, an identity statement of the form $a = a$ will fail to be true in every world, and so $\models_{prpr} a = a$ will not be the case. There are at least three responses that can be given to this. The first two both bite the bullet, in a sense, and accept the criticism (albeit differently), while the third accepts the criticism as compelling enough to force a re-evaluation of *prpr*.

The first response is perhaps the most dismissive. We might refer to it as a negative response. It is to say that *prpr* needn't concern itself with self-identity because pragmatic logics generally needn't concern themselves with it, nor indeed any similarly premise-less arguments, because if nothing has been said - that the information with which you are basing an inference is zero - then, pragmatically speaking, there is nothing with which to base one's conclusions. Another way to put this is that a pragmatic logic needn't concern itself with the *a priori*. In short, then, this response views the whole issue of self-identity as a moot point.[2] Although this position might have

[1] [45], p90e.

[2] I am reminded of a discussion I took part in when I was first being introduced to the basics of logic. My initial query lay in whether or not we should consider premise-less arguments sound or not. It wasn't clear to me whether the definition "a valid argument with true premises" required that there actually be premises, or just that any premises it had needed to be true. In any case, I thought, can't I start an argument with a logical truth? The premise would be redundant in a sense, but why couldn't I argue p or not p, therefore...? My professor at the time, who, it should be noted was of a legal bent, was very much against the idea, insisting that it would be pointless to construct such

some merit, it largely goes against the spirit of this paper. In forming *prpr* we were, after all, thoroughly concerned with what assumptions we should make prior to any *a posteriori* information; Default assumptions were seen to be very important for *prpr*. Whatever value is in the negative response, then, for the reason outlined, it cannot be a serious position (for us).

The second response is to accept the failure of self-identity, but to do so on the grounds that it says something philosophically important about the nature of vagueness. That in other words, self-identity *must* fail for some reason. We might refer to this as the positive response to the failure of self-identity, as opposed to the negative one that we just mentioned. As for what specific reason we might have, it is difficult to extend beyond mere speculation, but there may well be rational grounds for disallowing $a = a$ for all objects, vague or not. Consider, for example, the following utterance that you have no doubt heard (or said) before: "I don't feel like myself right now..." Perhaps the most comfortable way to interpret this is simply that the person means that they are not feeling like the person they felt previously, yesterday, say. But what if we were to take the utterance more literally? Could they really mean that this very instant they feel as though they are not themselves? What could this mean? One plausible way to make sense of this might be through linguistic vagueness. Consider the following highly idealized thought experiment. You have but two phrases with which to describe your whole world: "light" and "not light". You are sitting at the beach at the early evening of a midsummer' day. Slowly but surely the sun descends, and ever so gradually day becomes night. By hypothesis, you can't describe all of this, of course. Suppose, for the sake of argument, you can't even think it: your thoughts are wholly constrained by your language. Even so, it seems plausible that you might be able to think or say "light" and "no light" at the right moments, namely towards the beginning of the night and the end of it. But in that borderline between, the idea would be that your language (and so your mind) would fail you. You would not be able to identify this half-light state, it would be left as a blur: an indescribable, unintelligible absence of a definite self-identical object. Whether this really fits with the

arguments. He added that it would never find a place in court. We already know $p \vee \neg p$, so such arguments are pointless. This is roughly the idea I'm getting at in the negative response. The negative response views self-identity and other premise-less arguments in the same way as my teacher viewed the soundness of trivial truths. I'm still not entirely convinced on the soundness matter. Even in court I could imagine a lawyer using such an argument to force a reluctant witness to speak, saying, "You were either friends with my client or not, Mr. X. Which one was it?". Whether this is an argument or not, I'm not sure. Perhaps the issue is more about whether *everyone* knows trivial laws like p or $\neg p$.

dialethic reading of paraconsistency in *prpr* is questionable. It is also especially troubling with regards to how *prpr* deals with the sorites argument: before we can come to find inconsistency, every 'object' in the series is an indefinite blur - quite a paradox! Perhaps it is more understandable what emotions the person uttering the claim of non-self-identity must be feeling, then. This response would clearly need , but it does seem like a worthwhile direction in which to travel, if only because rejecting it would deny *prpr* of its ability to always demand consistency of identity statements.

If neither of these options are compelling, and self-identity is seen to be a principle that must be kept, we might withdraw towards the LP approach, which arguably gets things slightly better by at least allowing $a = a$. That is, even though inconsistency will demand that $a = a \wedge a \neq a$, it is still the case that $\models_{LP} a = a$. Self-identity is, for LP, a soft constraint. A way in which we might adapt *prpr* would be to take mimic what we did in ST and favour a mixed consequence relation, for example, *prt*. This would allow self-identity to be a law in the LP sense. This would also solve our problem regarding the fact that $a = b, a \neq b \not\models_{prpr} a = b \wedge a \neq b$, as $a = b, a \neq b \models_{prt} a = b \wedge a \neq b$, because with inconsistent premises, the pragmatic interpretation collapses to the tolerant one. *prt* has other nice features, as well, like (SI), (TI), (NCO) and (NVI). We would lose (ToI) though. A detailed look at *prt* and the version of the sorites considered would be worth pursuing. For some other problems this has been done (see [11]).

Another issue concerns our definition of identity in *prpr*. We argued that, although it did not concur with the perhaps more standard biconditional approach as seen in LP and ST, it still satisfied Leibniz's Law in some sense. Specifically, it satisfied the definition that $a = b$ should hold iff for every property, P, Pa and Pb would have the same truth-value. Even so, it might be said that the definition was ad hoc, and that it has drastic consequences for the account, namely that the truth-value definition no longer tallies with the biconditional definition. Interestingly, it might be said that the real problem lies with the way conjunction is defined in *prpr*. This is because conditionals by themselves appear to behave perfectly well. For example, unlike in LP, modus ponens holds in *prpr*. More generally, conditionals by themselves are perfectly well behaved consistent entities. In actual fact, during the process of coming to *prpr* as it is now, the idea of better capturing the biconditional by changing the definition was considered. The hope then, was that we could define identity in the biconditional way. In the end, the change, as we made it, led to other, less appealing properties. For example, premises would depend dramatically on their order, leading to some unwanted results. There may be some better way to do it, though, a way that both keeps the biconditional definition and leads to more reasonable results. This would have to be an issue for further research.

Aside from problems related to *prpr* as a whole, it might also be useful to look, for a final time, at our treatment of identity across all three of our logics.

5.2 A Final Look at Identity

Arguably, all our logics deal with a changing object in the same way: by resorting to a third truth-value (whether this be distinct from true and false, or just a combination of them). In the face of many challenges, then, all three approaches will live or die by the same sword. Perhaps the most common criticism offered is that the addition of a third, or indeed, an $n-th$ value does not help with matters of vagueness, and so also our approach with changing objects. This is summed up well by Sainsbury's "You do not improve a bad idea by iterating it" ([39] p. 255). The criticism is usually that whereas before we had only one vagueness problem, i.e. between the certainties of truth and falsity, by adding a third value we now have two: the vagueness between strict truth and borderline, and borderline and strict falsity. We can keep on adding values, but we'll just end up with more and more vagueness. This would have to stop, at some point, we would think. At some point, we might put it that there would be a sufficient number of truth-degrees for total meta-vagueness to be present. That we would have, perhaps, a continuum of values, a fuzzy logic. Although this needn't be any better than what we have, it does highlight the perhaps obvious fact that if we never could improve a bad idea by iterating it, then any field of study having connection to statistics or the like would be fatally flawed. This view seems to suggest that all our logics fall short. They are three-valued and no more, and this "no more" means that they don't iterate the bad idea enough. While this process would be easy for LP and ST given the formalism in chapter three, the notation in chapter four would make it a little harder for *prpr* at present. In principle it should be possible, though, and this approach might be worth investigating.

Another communal problem that all our approaches share is that they don't provide any resolution to what we called the puzzle (as opposed to the paradox) of the sorites. This was the problem of where to place the borderline. Traditionally, "borderline" here could perhaps mean the abstract non-position that lies between truth and falsity - or better put, in the case for identity, the position 'between' some predicate that holds and predicate that doesn't. The less abstract meaning of borderline would be the last property before the change, i.e. there would be two borderlines. Where we are using the three-values of our logics, this can, in some cases, be uniquely identifiable, for example, as with the case of the 'bridge' amoeba in Prior's example. In any case, the point is that none of our logics tell us where borderlines lie. Perhaps this is as it should be. Perhaps such things are empirical matters, and our logics, then, should be viewed as flexible in that they can deal with a wide range of data. It does seem plausible though, that borderlines be constant, in the sense that we always found them in the same places. Just how constant would be an interesting question to look at. Where borderlines be may depend on what predicate one is considering, perhaps, or one's physiology, but it seems possible that this could all be factored into a logic. We could then plug in the information and the logic would give us answers to the puzzle of the sorites, as well as the paradox. Again, this would suggest a "more values" many-valued approach, if we wanted to keep any of the logics in this paper.

In conclusion, our approach was to consider three logics that all could all deal with sorites-like change while holding on to a Leibnizian account of identity. With LP, we saw how this was achieved through the non-transitivity of identity, which could allow for borderline cases where predicates could be said to both hold and not hold for a given object. The account of identity was strong enough to account for the premise that the change was gradual so as to allow identity between objects at each successive step. We saw that the problematic conclusion was not reachable in light of this non-transitivity. In ST, we kept (TI), but saw how the non-transitivity of ST inference would not allow the chaining together of otherwise valid arguments. We saw that, because ST allowed for (SI), we could move one further step along in compared to LP, but then the argument would break off. Despite some of the admirable properties of ST, for example, that it doesn't conclude the problematic conclusion, and that it holds (TI) to be true, we found that, as presented, this version of ST was not capable of truly capturing the spirit of a paraconsistent logic. We then saw how a logic could be constructed by inspiration of pragmatic considerations regarding vagueness. Specifically, we argued that *prpr* was plausibly more viable than LP because it would allow the general form of the sorites argument to go through, which we should expect, but that with new information of inconsistency, due to the non-monotonic nature of the logic, we could reinterpret our premises to allow for borderline objects. The result was such that all our objects would become borderline, which we argued might explain the troubling aspect of the sorites paradox. Again, with new information, *prpr* could make it so that we fall back into a more *LP*-style reading of the sorites, so that there would be minimal borderlines. For these reasons, together with its successful treatment of all the target features, we said that *prpr* was the most viable logic of the three we considered, taking in a sense, the best of both aspects from the *LP* and *ST* accounts. We went on to consider possible problems in the account, including the case of self-identity, and we also considered the adoption of another logic with beneficial properties, namely *prt*. If there was an overall message to describe what we have found in this paper, it would be that while paraconsistency can help, it needn't overshadow consistency, and that we can take the most agreeable aspects of both to make sense of the nature of change.

Appendix - *prpr* in Python

The following code (which was written in Python 3 but should also run in Python 2) gives a sense of *prpr* in action. The user is prompted to input some predicates and objects, both of which must be given in the form of a list of strings (e.g. [“F”, “G”] and [“a”, “b”]). The program will generate the set (- really a list -) of possible worlds, and the user can then use the defined functions to perform various logical operations. The most important of these functions is probably prpr, which takes as arguments a list of premises and a conclusion and returns *True* if the argument holds in *prpr* and false otherwise (e.g. prpr([“Fa”, neg(“Fa”)], “Fa”) returns False, as we should expect). See code comments for more details. The script is neither optimized for efficiency, nor has it been tested to work for every possible sentence form. The number of predicates and variables should be kept to a minimum to ensure results within a reasonable time (less than 5 in total, should be fine).

```
import itertools
from itertools import chain

predicates = input("What predicates do you want?\n")
domain = input("What objects?\n")

def opposite(x):
    if x[-1] == "r":
        return x[0]+x[1]
    else:
        return x + "-bar"

## creation of state of affairs  ##
def SOA():
    result = []
    for i in predicates:
        for j in domain:

            result.append(i+j)
            result.append(opposite(i+j))
    return result

def positive_SOA():
    result = []
    for i in predicates:
        for j in domain:
            result.append(i+j)
    return result

def facts():
    result = [[]]
    for i in positive_SOA():
        temp = [x for x in result]
        result = []
        for j in temp:
```

```
##              result.append([i])
            will_pos = [ooh for ooh in j]
            will_pos.append(i)
            result.append(will_pos)

            will_pos = [ooh for ooh in j]
            will_pos.append(opposite(i))
            result.append(will_pos)

            will_pos = [ooh for ooh in j]
            will_pos.append(i)
            will_pos.append(opposite(i))
            result.append(will_pos)

    return result


####    SET OPERATIONS    #############
def union(x,y):
    result = []
    for i in x:
        if i not in result:
            result.append(i)
    for j in y:
        if j not in result:
            result.append(j)
    return result

def intersection(set_1, set_2):
    result = []
    for i in set_1:
        if i in set_2:
            result.append(i)
    return result

def ocross(x,y):
    temp = []
    for A in x:
        for B in y:
            wim = union(A,B)
            temp.append(wim)

    return union(temp,temp)

##   NEGATION   ####################
def neg(x):
    return ["neg", x]

## DISJUNCTION ##########

def disj(x,y):
    return ["disj", x, y]

##   CONJUNCTION #######################
# superficial construction ##
def conj(x,y):
    return ["conj", x, y]

def excl(p):
    return [p, opposite(p)]

## no new inconsistency test ##
def nnct(set_1, set_2, union_set):
    result = True
    for p in positive_SOA():
        if Set(excl(p)).issubset(Set(union_set)):
            if Set(excl(p)).issubset(Set(set_1)) == False:
                if Set(excl(p)).issubset(Set(set_2)) == False:
                    return False
    return True

def T_conj(x,y):
    return ocross(x,y)


##   DISJUNCTION #####################
def T_disj(x,y):
    return union(x,y)
#( also F_conj) #

def imp(x,y):
    return ["imp", x, y]
```

```
## UNIVERSAL QUANTIFICATION ######
def forall(variable, formula):
    return ["forall", variable, formula]

def T_forall(variable, formula):
    result = list(formula)
    temp = [i for i in result]
    berr = [i for i in temp]
    for i in range(0, len(domain)):
        print i
        for j in range(0, len(formula)):

            text = formula[j]
            new = list(text)
            tutem = new
            print "tutem= ", tutem
            for k in range(0, len(new)):
                if new[k] == variable:
                    tutem[k] = domain[i]
            temp[j] = ''.join(tutem)
        if i == 0:
            barb = T(temp)
        else:
            barb = ocross(barb, T(temp))
        print "temp= ", temp, "barb= ", barb

    return barb

## SIMILARITY RELATION:: ##############

def sim(predicate,a,b):
    return ["sim",predicate, a,b]

def T_sim(predicate, a, b):
    return [[predicate+a[0], predicate + b[0]],[predicate + a[0] + "-bar", predicate + b[0] + "-bar"]]

def F_sim(predicate, a, b):
    return [[predicate+a[0], predicate + b[0]+"-bar"],[predicate + a[0] + "-bar", predicate + b[0] ]]

## IDENTITITY::    ##################
def equal(a,b):
    return ["equal", a, b]

def T_equal(a,b):
    result = T(sim(predicates[0], a, b))
    for i in range(1, len(predicates)):

        temp = T(sim(predicates[i], a, b))
        result = ocross(result,temp)
    return result

def F_equal(a,b):
    result = F(sim(predicates[0], a, b))
    for i in range(1, len(predicates)):

        temp = F(sim(predicates[i], a, b))
        result = union(result,temp)
    return result

def F_conj(a,b):
    return union(F(a),F(b))

## TRUTH-MAKERS AND FALSE-MAKERS CONJUNCTIONS ##########
def T(phi):
    if phi in SOA():
        return [[phi]]

    if phi[0] == "neg":
        return F(phi[1])

    if phi[0] == "disj":
        t1 = phi[1]
        t2 = phi[2]
        return T_disj(T(t1), T(t2))

    if phi[0] == "conj":
        t1 = phi[1]
        t2 = phi[2]
        return T_conj(T(t1),T(t2))

    if phi[0] == "imp":
        return T(neg(conj(phi[1],neg(phi[2]))))
```

```
    if phi[0] == "forall":
        t1 = phi[1]
        t2 = phi[2]
        return T_forall(t1, t2)

    if phi[0] == "sim":
        t1 = phi[1]
        t2 = phi[2]
        t3 = phi[3]
        return T_sim(t1,t2,t3)

    if phi[0] == "equal":
        t1 = phi[1]
        t2 = phi[2]
        return T_equal(t1,t2)

def F(phi):
    if phi in SOA():
        if phi[-1] == "r":
            return [[phi[0]]]
        else:
            return [[phi+"-bar"]]

    if phi[0] == "neg":
        return T(phi[1])

    if phi[0] == "disj":
        t1 = phi[1]
        t2 = phi[2]
        return T_conj(F(t1), F(t2))

    if phi[0] == "conj":
        t1 = phi[1]
        t2 = phi[2]
        return T_disj(F(t1),F(t2))

    if phi[0] == "imp":
        return T(conj(phi[1],neg(phi[2])))

    if phi[0] == "sim":
        t1 = phi[1]
        t2 = phi[2]
        t3 = phi[3]
        return F_sim(t1,t2,t3)

    if phi[0] == "equal":
        t1 = phi[1]
        t2 = phi[2]
        return F_equal(t1,t2)

W = facts()

from sets import Set

def consistent_worlds(world):
    temp = []
    for i in SOA():
        if i in world and opposite(i) not in world:
            if i not in temp:
                temp.append(i)
    return temp

def inconsistent_worlds(world):
    temp = []
    for i in SOA():
        if i in world and opposite(i) in world:
            if i not in temp:
                temp.append(i)
    return temp

def less_than(f, w, v):
    temp_1 = []
    temp_2 = []
    for i in SOA():
        if i in f and opposite(i) in v:
            temp_1.append(i)
        if i in f and opposite(i) in w:
            temp_2.append(i)
    check = True
    for i in temp_1:
        if i not in temp_2:
            check = False
    if check == True and temp_1 != temp_2:
```

```
        return True
    else:
        return False

def toler_worlds(phi):
    result = []
    for w in W:
        for f in T(phi):
            if Set(f).issubset(Set(w)):
                result.append(w)
    return result

def tol_worlds_Gamma(Gamma):
    sclemp = toler_worlds(Gamma[0])
    for gamma in Gamma:
        scosclemp = []
        for i in sclemp:
            if i in toler_worlds(gamma):
                scosclemp.append(i)
        sclemp = scosclemp
    toler_Gamma = sclemp
    return toler_Gamma

def PRAG(phi, worlds):
    result = []
    for w in worlds:
        check = False
        for f in T(phi):
            if Set(f).issubset(Set(w)):
                check = True
                check_x = True
                while check_x == True:

                    for v in worlds:
                        if Set(f).issubset(Set(v)):
                            if less_than(f,w,v) == True:
                                check = False
                                check_x = False
                    check_x = False
        if check:
            result.append(w)
    return result

def prpr(Gamma, phi):
    temp = []
    if Gamma == []:
        if PRAG(phi, W) == W:
            return True
        else:
            return False
    else:
        temp = PRAG(Gamma[0],W)
        for gamma in Gamma:
            sec_temp = PRAG(gamma, W)
            thir_temp = []
            for i in temp:
                if i in sec_temp:
                    thir_temp.append(i)
            temp = thir_temp
        if temp != []:
            for i in temp:
                if i not in PRAG(phi, W):
                    print temp
                    return False
            print temp
            return True
        else:
            temp = PRAG(Gamma[0],tol_worlds_Gamma(Gamma))
            for gamma in Gamma:
                sec_temp = PRAG(gamma, tol_worlds_Gamma(Gamma))
                thir_temp = []
                for i in temp:
                    if i in sec_temp:
                        thir_temp.append(i)
                temp = thir_temp
            for i in temp:
                if i not in PRAG(phi, tol_worlds_Gamma([phi])):
                    print temp
                    return False
            return True

## checks substitutivity of identicals
def check_SI():
```

```
    return prpr(["Fa",equal("a","b")],"Fb")
## checks transitivity of identicals
def check_TI():
    return prpr([equal("a","b"),equal("b","c")], equal("a","c"))
```

Bibliography

[1] Alxatib, S., Pelletier, F.R. (2011). The psychology of vagueness: borderline cases and contradictions. Mind and Language, 26, 287326.

[2] Aristotle, (1963), 'Aristotle in twenty-three volumes', The Loeb Classical Library, London: Heinemann, Harvard University Press.

[3] Asenjo, F., G., (1966), 'A Calculus of Antinomies', Notre Dame Journal of Formal Logic, 7: 1035.

[4] Barrio, E., Rosenblatt, L., Tajer, D., 'The Logics of Strict-Tolerant Logic', Journal of Philosophical Logic, October 2015, Volume 44, Issue 5, pp 551571.

[5] Max Black, "Identity of Indiscernibles," Mind, vol. LXI, no. 242, 1952.

[6] Cobreros, P., Egre, P., Ripley, D., and van Rooij, R. (2012). Tolerant, classical, strict. Journal of Philosophical Logic, 41(2):347-385.

[7] Cobreros, P., Egre, P., Ripley, D., and van Rooij, R. (2013), 'Reaching Transparent Truth', Mind, Vol. 122 . 488 .

[8] Cobreros P., Egré P., Ripley D., van Rooij R., (2013) *Identity, Leibniz's Law and Non-transitive Reasoning*, Metaphysica, 2013, Vol.14(2), pp.253-264.

[9] Egr, P., V. de Cardelle, D. Ripley (2013). Vagueness and order effects in color categorization. Journal of Logic, Language and Information, 265300.

[10] Cobreros, P., Egre, P., Ripley, D., and van Rooij, R. (2014), Priests' Motorbike and Tolerant Identity', (2014), in H. Wansing et al (eds.), Recent Trends in Philosophical Logic , Springer, pp. 75-84.

[11] Cobreros, P., Egré, P., Ripley, D., van Rooij R., (2015), 'Pragmatic interpretations of vague expressions: strongest meaning and nonmonotonic consequence', Journal of Philosophical Logic , 2015, 44 : 275-393.

[12] Cobreros P., Egré P., Ripley D., van Rooij R., (2016) *Comparing Some Substructural Strategies Dealing with Vagueness* Information Processing and Management of Uncertainty in Knowledge-Based Systems: 16th International Conference, IPMU 2016, Eindhoven, The Netherlands, June 20 - 24, 2016, Proceedings, Part II, pp. 161-172, Springer International Publishing.

[13] Evans, G., 'Can There Be Vague Objects?', Analysis, Vol. 38, No. 4 (Oct., 1978), p. 208.

[14] Frege. G., (1892), On Sense and Reference, in P. Geach and M. Black (eds.) Translations from the Philosophical Writings of Gottlob Frege, Oxford: Blackwell (1952).

[15] Grice, H.P. (1967), *Logic and conversation*, William James Lectures, Harvard University Press, reprinted in Studies in the Way of Words, 1989, Harvard University Press, Cambridge, Massachusetts.

[16] Hacking, I., (1975), "The Identity of Indiscernibles", Journal of Philosophy, 72 (9): 249-256.

[17] Haack, S., (1978), 'Philosophy of Logics', Cambridge University Press.

[18] D. Hyde (1997), From Heaps and Gaps to Heaps of Gluts, Mind 106, 641-60.

[19] Kaufman, E.L., Lord, M.W., Reese, T.W., Volkmann, J. (1949). 'The discrimination of visual number', The American Journal of Psychology. 62 (4): 498525.

[20] Kleene, S., C., 'On a notation for ordinal numbers', The Journal of Symbolic Logic 3 (1938), pp. 150155.

[21] Łukasiewicz (1920), 'O logice trjwartociowej' (in Polish), Ruch filozoficzny 5:170171. English translation: On three-valued logic, in L. Borkowski (ed.), Selected works by Jan ukasiewicz, NorthHolland, Amsterdam, 1970, pp. 8788.

[22] Magidor O. (2011) Arguments by Leibniz's Law. Philosophy Compass 6: 180195.

[23] Mates, B., 'Elementary Logic', second edition, Oxford University Press.

[24] Post, E., L., (1943), 'Formal Reductions of the General Combinatorial Decision Problem', American Journal of Mathematics. 65: 197215.

[25] Priest, G., G., (1979), 'The Logic of Paradox', Journal of Philosophical Logic 8, pp. 219-241.

[26] Priest, G., G., (1991), 'Sorites and Identity', Logique Analyse, 135-136, pp.293-296.

[27] Priest, G., G., (1991) 'Minimally Inconsistent *LP*', Studia Logica: An International Journal for Symbolic Logic, 1991, 50 (2) pp.313-331.

[28] Priest, G., G., (1997) 'Inconsistent Models of Arithmetic Part I: Finite Models', Journal of Philosophical Logic, 26 (2) 223235.

[29] Priest, G., G., (2002), 'The hooded man', Journal of Philosophical Logic, 31, pp. 445-467.

[30] Priest, G., G., (2008), 'An Introduction to Non-Classical Logic: From If to Is', Cambridge: University of Cambridge Press.

[31] Priest, G., G., (2010), 'Non-Transitive Identity'. in R Dietz S Moruzzi (eds), Cuts and Clouds: Vagueness, its nature and its Logic. Oxford University Press, pp. 406-416.

[32] Priest, G., G., (2014), 'One: being an investigation into the unity of reality and of its parts, including the singular object which is nothigness', Oxford University Press, 2014.

[33] Prior, A. N., (1968), 'Papers on Time and Tense', Oxford Clarendon press.

[34] Quine, W.V.O., (1961), 'From a Logical Point of View', Cambridge MA: MIT Press.

[35] Read, S. 1995: Thinking About Logic: An Introduction to the Philosophy of Logic. Oxford: Oxford University Press.

[36] Ripley, D., (2011), 'Contradictions at the borders', in R. Nouwen, R. van Rooij, H.-C. Schmitz, U. Sauerland (Eds.), (2012), Vagueness in communication. LICS: Springer.

[37] Rips, L. J., (2011), 'Split identity: Intransitive judgments of the identity of objects', Cognition 119: 356373.

[38] Sainsbury, R. M., 1987: 'Paradoxes', Cambridge: Cambridge University Press, 1995. Originally published in 1987.

[39] Sainsbury, R. M., (1996), Concepts without boundaries, Vagueness: A Reader, in Keefe and Smith (eds.) 1996, 251264.

[40] Serchuk, P., Hargreaves, I., Zach, R. (2011), *Vagueness, logic and use: Four experimental studies on vagueness*, Mind and Language, 26: 540-573.

[41] Smiley, T.J., 1960, 'Sense without denotation', Analysis 20.

[42] van Fraassen, B. C., (1969), 'Facts and Tautological Entailment', The Journal of Philosophy, vol.66, no.15, pp.477-487.

[43] Williamson, T., Vagueness in Reality M. Loux and D. Zimmerman, eds., The Oxford Handbook of Metaphysics, Oxford: Oxford University Press.

[44] Williamson, T., (1994) 'Vagueness', London: Routledge.

[45] Wittgenstein, L., 'Philosophical Investigations', 4th edition, 2009, P.M.S. Hacker and Joachim Schulte (eds. and trans.), Oxford: Wiley-Blackwell.