Essays on Identity: A Defense of Logical Orthodoxy

by

Sergio Gallegos

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

2011

This manuscript has been read and accepted for the Graduate Faculty in Philosophy in satisfaction of the dissertation requirements for the degree of Doctor of Philosophy.

Barbara Montero

_____ _____

Date Chair of Examining Committee

Iakovos Vasiliou

_____ _____

Date Executive Officer

Richard Mendelsohn

_____

Arnold Koslow

_____

Alberto Cordero

_____

Supervisory committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Essays on Identity: A Defense of Logical Orthodoxy

by

Sergio Gallegos

Advisor: Professor Richard Mendelsohn

My dissertation defends a commonly accepted package involving a certain number of theses that lie at the intersection of metaphysics, philosophy of language and philosophy of logic. The package, which includes (i) the classical thesis that identity is a one-one relation that is absolute, necessary and determinate and (ii) the Kripkean theses that true identity statements involving only rigid designators are necessary and that proper names are rigid designators (along with the consequences of these two theses such as the view that true identity statements involving only proper names are necessary), occupies a central role in many philosophical discussions where it functions both as a sanction of certain avenues of inquiry and as a constraint on the development of others.

In spite of its inner consistency and overwhelming persuasiveness, the package has been criticized from many different angles. For instance, some philosophers claim that there are many different identity relations relativized to different sorts of things rather than an absolute identity relation because the latter view is undermined by paradoxes. Others maintain that, *pace* Kripke, there are true contingent identity statements involving only proper names because the reference of a proper name in a counterfactual situa-

tion is given by a sortal concept associated to it. Furthermore, some hold that identity is indeterminate in certain circumstances because the strongest arguments that aim to show the inconsistency of the view that identity is indeterminate may be blocked successfully.

The first chapter of my dissertation is concerned with answering the challenge raised against the absolute character of identity. After reviewing the traditional considerations put forward by relative identity theorists as well as some novel arguments, I conclude that identity is absolute. In the second chapter, I consider in detail some of the most prominent arguments given to maintain that there are true contingent identity statements involving proper names as well as an argument given to show that the proof of the necessity of identity involves a vicious circularity and I show that all the arguments involve serious flaws, thus clearing of doubts the Kripkean portion of the package. The third chapter vindicates the view that indeterminate identity is inconsistent by providing a defense of Evans' argument for the inconsistency of indeterminate identity against a number of objections that have been addressed to it.

Acknowledgments

First of all, I wish to thank Richard Mendelsohn for being such a great advisor. For some graduate students, writing a dissertation can become a tortuous and neverending process where you constantly question your capacity for articulating original thoughts and your mental endurance. Richard provided the required support and guidance to help me through the most difficult stages of writing, never losing his patience and always providing me with sound advice. After I took the decision to move temporarily to Venezuela with my familiy to write the bulk of the dissertation, he supported the move and continued to provide valuable suggestions and feedback via e-mail. I wish to thank him in particular for always reminding me that the dissertation is not the end of one's professional career, but only the beginning.

I also wish to thank Arnold Koslow and Alberto Cordero. Both have been among the most important influences in my graduate school years. Both have been, not only sources of inspiration and intellectual models to be followed, but constants reminders —through their words and actions— of the main reasons that led me to pursue a career in philosophy. My graduate school years would have been far more difficult and far less enriching without their constant encouragement and advice. In addition, I want to express my gratitude to both Barbara Montero and John Greenwood who read the dissertation on a very short notice and made some very interesting and valuable suggestions concerning some portions of my project during the defense.

I want to acknowledge the financial support of the National Council for Science and Technology of Mexico (CONACYT) and the Mexican Department of Public Education (SEP). These two institutions deserve special men-

tion for providing me with partial support during the initial stages of my graduate studies. In addition, I want to acnknowledge the financial support of CUNY through a Writing Fellowship that helped me carry out the initial stages of the research for the dissertation.

I want to thank a number of remarkable women without whom this disseration would have never been completed. Lourdes Valdivia suggested the topic of this dissertation. Her keen philosophical acumen helped me avoid numerous pitfalls. She has been, as much as Richard, Arnie and Alberto, an intellectual model to be followed and a dear friend. My mother Martha Ordorica supported me emotionally and financially during my graduate school years with unwavering love and affection. My sister Paulina Gallegos helped me get through some of my most serious crises of faith with her characteristic no-nonsense attitude and her constant affection. My wife Alejandra Myerston, who has patiently endured being married to a philosopher, deserves special recognition for believing in me in the most trying times, for putting up with the lifesyle and the demands that my graduate career imposed on us and for allowing me to be a part of her life. One cannot wish a better companion than her to sail the seas of time.

I dedicate this dissertation with love to my father Sergio Gallegos, who passed away before its completion, and to my daughters Aurelia and Mikaela.

S.G.

May of 2011

New York, NY

Contents

1 Introduction **1**

2 Rejecting relative identity **10**

2.1 Introduction 10

2.2 RIT: the main theses 13

2.2.1 The non-existence of classical identity 14

2.2.2 The sortal relativity of identity 19

2.2.3 The derelativization thesis 28

2.3 Geach's arguments for RIT 34

2.3.1 Geach's arguments for RIT_1 34

2.3.2 Geach's arguments for RIT_2 46

2.3.3 Geach's argument for RIT_3 54

2.4 Some traditional objections against RIT 56

2.4.1 Some traditional replies against RIT_1 57

2.4.2 Some traditional replies to RIT_2 67

2.4.3 Some traditional replies to RIT_3 79

2.5 Some novel arguments in favor or RIT 81

2.6 Conclusion 86

3 Vindicating the necessity of identity **88**

3.1 Introduction 88

3.2 Gibbard's defense of contingent identity 95

3.2.1 Kripke's argument for the rigidity of names 95

3.2.2 The Goliath-Lumpl scenario and Gibbard's theory of names 98

3.2.3 A refutation of Gibbard's views 116

3.3 Wreen's challenge to the necessity of true identity statements involving names 137

3.3.1 Wreen's argument 138

3.3.2 Some problems with Wreen's argument: inconsistent assumptions and conflicting interpretations 141

3.3.3 Wreen's argument and Kaplan's distinction between contexts and circumstances of evaluation 146

3.4 Della Rocca's objection to the necessity of identity 160

3.4.1 Della Rocca's objection 161

3.4.2 Essentialism and the necessity of identity 166

3.4.3 A justification of Kripke's discriminatory treament of modal intuitions 171

3.5 Conclusion 178

4 The determinacy of identity **179**

4.1 Introduction 179

4.2 Evans' argument 187

4.3 Some interpretive matters 192

4.4 Two objections from Parsons and Woodruff 199

4.4.1 Some background assumptions 200
4.4.2 The two objections 203
4.4.3 Responses to the objections 209
4.5 Van Inwagen's objection 221
4.5.1 The background discussion: the special composition question 222
4.5.2 The objection: Van Inwagen's semantics as a tool to block the *reductio* 227
4.5.3 Some difficulties with Van Inwagen's proposal 233
4.6 Williams' response to the *reductio* 242
4.6.1 The loophole in Evans' *reductio* 243
4.6.2 Williams' modal framework: multiple actualities 246
4.6.3 Some difficulties with Williams' response 252
4.7 Barnes' objection 259
4.7.1 Barnes' framework of modal vagueness 259
4.7.2 The counterpart-theoretic interpretation of Evans' *reductio*: motivations and implications 262
4.7.3 A critical assessment of Barnes' response 269
4.8 Conclusion 277
Bibliography 279

Chapter 1

Introduction

'Identity', according to David Lewis (1986: 192), 'is utterly simple and unproblematic.' Lewis elaborates on this assertion about the simple and unproblematic nature of the identity relation by stating that 'there is never a problem about what makes something identical to itself; nothing can ever fail to be. And there is never a problem about what makes two things identical; two things can never be identical.' These remarks, which echo Wittgenstein's observation in the *Tractatus* (1922: 105) that 'to say of *two* things that they are identical is nonsense and to say of *one* thing that it is identical with itself is to say nothing at all', suggest that philosophical discussions about identity are in a sense pointless because identity is a trifling and uninteresting relation. Indeed, given that everything is identical to itself, the relation of identity, which can be characterized as the only equivalence relation that satisfies the Indiscernibility of the Identical, holds trivially of every object and itself and thus fails to characterize or demarcate in any relevant way any object or any category of objects.

However, though many philosophers acknowledge the simple and unproblematic nature of identity, there seems to be a significant number of philosophical discussions that are concerned with the notion. For instance, there is a persisting controversy concerning the conditions under which some person P that exists at a time t can be rightly said to be identical to some person P' that exists at a different time t'. The notion of identity also appears to be involved in the debate concerning the issue of whether it is legitimate to talk about the on-goings of some actual object o in different counterfactual circumstances as on-goings of *that* object. In addition, there is also a discussion (which is motivated by thought experiments that are variations of the traditional Ship of Theseus puzzle) concerning whether it makes sense to maintain that some object o is indeterminately identical to some object o' in certain cases.

Lewis' take on these discussions, which is shared by many other philosophers (e.g., Jubien 1996 and McGinn 2000), is that, even though they are concerned with important philosophical problems, these problems are not problems about the identity relation itself, but rather problems about other notions formulated in terms of identity. For instance, the controversy concerning the nature of the conditions under which some person P that exists at a time t can be rightly said to be identical to some person P' that exists at a time t' can be reformulated as a controversy concerning the notion of person in the following way: is a person an entity such that it can exist at different times? Similarly, the debate concerning the legitimacy of talking about the on-goings of an actual object o in counterfactual circumstances as on-goings of *that* object can be reformulated according to Lewis as a mere-

ological question: is o an object such that it can be both part of actual circumstances and part of certain counterfactual circumstances?

The strategy suggested by Lewis to handle the aforementioned 'identity problems' is very appealing insofar as it shows that the problems in question are genuine philosophical problems that deserve attention but that they do not really undermine the intuitive thesis that identity is a simple and unproblematic relation. However, even if we endorse Lewis' viewpoint, there are some worrisome questions that persist. For instance, how can identity be a simple and unproblematic relation if it holds between every object and itself (i.e., if it holds across all ontological categories)? The intution behind this worry is that, if we accept that any relation R must be defined over a certain domain of entities $\mathcal{D}$ (e.g., numbers or physical objects) and that the relation in question is not defined over another domain of entities $\mathcal{D}'$ (which is the co-domain of $\mathcal{D}$), identity is not a genuine relation because it is defined over an all-encompassing domain that does not exclude anything and any reference to such a domain generates paradoxes. Some authors such as Geach (1967) rely on this intuition to argue that the thesis that Lewis endorses concerning identity (and that the vast majority of philosophers share) rests on a false assumption —namely, the assumption that there is only one identity relation that holds between every object and itself regardless of the nature of the object. Geach has further argued that, once this assumption is rejected, one may show that identity is relative to different categories of entities which fall under different sortals.

In addition to the aforementioned worry concerning the fact that identity is a simple and unproblematic relation defined over an all-encompassing

or universal domain, there are other troublesome doubts that threaten to undermine the Lewisian view (which I will refer to subsequently as the 'orthodox' view of identity). For instance, though Marcus (1947) proved that the identity relation is necessary and Kripke (1980) defended this conclusion by arguing persuasively that the intuition that there are contingent identities can be successfully explained away by distinguishing the identity relation from identity statements and admitting that some identity statements can be contingently true when they involve what Kripke calls 'non-rigid designators' while maintaining that all true identity statements that involve exlusively 'rigid designators' are necessary, some doubts persist concerning Kripke's attempt to explain away the intuition that there are contingent identities.

In particular, Gibbard (1975) has argued that Kripke's conclusions are too radical insofar as there are cases of contingent identity statements that involve exclusively names (which, according to Kripke, are paradigmatic rigid designators). Drawing on Gibbard's arguments, other authors have manifested related doubts vis-à-vis the Kripkean theses that identity is necessary and that all identity statements involving exclusively rigid designators are necessary. For instance, Della Rocca (2002) has contended recently that the proof of the necessity of identity that Kripke offers is flawed because it requires an essentialist assumption in order to be carried out but this essentialist assumption cannot be appropriately defended without relying on the necessity of identity —which makes the proof viciously circular.

Finally, there is another another source of doubt with respect to the orthodox view of identity. Indeed, even though both Evans (1978) and Salmon (1981) have argued persuasively that the identity relation is not vague or in-

determinate because, if it were the case that some object a is indeterminately identical to an object b then b would have a property that a lacks (namely, the property *being indeterminately identical to* a) and a would then be distinct from b instead of being indeterminately identical to it, both arguments have met opposition from several philosophers that have raised a number of problematic considerations concerning the arguments deployed by Evans and Salmon. In particular, Parsons and Woodruff (1995) have claimed that Evans' argument against the indeterminacy of identity is invalid because it appeals to the contrapositive of the Indiscernibility of the Identical (i.e., the Distinctness of the Discernibles) and the use of contrapositive principles is is illegitimate in a framework that allows ontic vagueness. In addition, Van Inwagen (1990) has maintained that Evans' argument against the indeterminacy of identity founders to the extent that the argument needs to be assesed in a non-classical logical framework in virtue of the resources that Evans introduces to construct it and that there are non-classical models in which the argument turns out to be invalid.

Thus, even though the orthodox view of identity appears to be intuitively correct, it faces certain doubts and potential objections that have motivated several authors to develop and endorse revisionist positions. To be more specific, the aformentioned considerations (as well as some others that I omit here for the sake of brevity) have fostered the development of positions that accept the existence of a multitude of relative identity relations and reject as nonsense the idea of an absolute identity relation or that tolerate the existence of cases where the identity relation (and not merely certain identity statements) is contingent or indeterminate. These revisionist positions are

problematic because the notion of identity occupies a fundamental role in our conceptual framework as it underlies many other important notions in semantics, logic and mathematics (e.g., the notion of synonymy, insofar as synonymy is taken to be *identity* in meaning, or the notion of set, insofar as our conceptual grasp of it depends on the axiom of extensionality which states that A and B are the same set if and only they have the *same* elements). Thus, adopting a revisionist position about identity involves an enormous cost because it forces us to rethink or even give up several central notions in logic and semantics as Hawthorne observes in the following passage:

> What, for example, is to count as an extensional context? What is it to mean to say that two terms corefer? All of these notions are built upon the notions of simple identity and difference. Abandon those notions and the intelligibility of a large range of logico-semantic concepts is cast into doubt. (Hawthorne 2006: 13-14)

In light of the huge cost that the adoption of any of the abovementioned revisionist positions entails, it would be desirable to respond to the main motivations put forward to defend them and, accordingly, to offer a vindication of the orthodox view of identity. This is the main purpose of this dissertation. I offer here a defense of the thesis that identity is a simple and unproblematic relation that holds exclusively between every object and itself (or, in other terms, that identity is the only equivalence relation that satisfies the Indiscernibility of the Identical) and I also defend the main consequences of this thesis —namely, the thesis that identity is necessary and the thesis that it is determinate. I do this by addressing some of the most prominent challenges to the orthodox view.

Though I intend to offer a vindication of the orthodox view by responding to some of the most important arguments and considerations that motivate the main revisionist positions, there are a number of considerations and positions that I will not consider in this dissertation. For instance, I will leave aside considerations regarding the apparent indistinguishability of an object and the sum of its constituent parts that some authors (e.g., Baxter 1988) rely on to argue for the revisionist thesis that the mereological relation of composition (which is a many-one relation that holds between a whole and its parts) is a form of identity. Even if I take this to be one of the most serious challenges to the orthodox view, a thorough refutation of the revisionist position that this challenge suggests would require a book-length treatment. Thus, I leave this for a future project.

Here is a roadmap of the dissertation. In the first chapter, I rehearse in detail Geach's relative identity theory, paying attention to distinguish thoroughly the different theses that make up his theory and to highlight the different relations between these theses, and I also present the main arguments that Geach puts forward to defend his theory —which he presents as alternative to the orthodox view. After showing that Geach's relative identity theory is essentially motivated by considerations stemming from the idea that admitting a relation such as identity that holds universally (i.e., across all ontological categories) forces us to be committed to an extravagant ontology and to accept paradoxes, I argue that Geach's attempts to show that advocates of the orthodox view are burdened by these problems ultimately founder in virtue of the fact that, in order to be able to formulate his objection, Geach is compelled to use the notion of absolute identity, which is part

of the orthodox view that he deems to be nonsense.

Having established that there are solid reasons to reject Geach's relative identity theory, I turn in the second chapter to defend two Kripkean theses which are consequences of the orthodox view: the thesis that identity is necessary and the thesis that true identity statements that involve exclusively rigid designators are necessary. After presenting a famous counterexample developed by Gibbard (1975) to Kripke's view about the modal status of true identity statements flanked by rigid designators and rehearsing an argument given by Wreen (1998) to argue that the identity statement 'Hesperus is Phosphorus' is a contingent truth, I show that Gibbard's counterexample fails to undermine Kripke's view on identity statements flanked by rigid designators because the semantics that Gibbard introduces to justify his claim makes clear that the alleged counterexample to Kripke's view is not even an identity statement by Gibbard's standards and I also show that Wreen's argument is flawed insofar as it fails to distinguish, as Kaplan (1989) does, contexts of utterance from circumstances of evaluation. Subsequently, I consider in the last section of this chapter an argument developed by Della Rocca (2001) to show that the proof of the necessity of identity is undermined by a vicious circularity, and I argue that this is not the case because Kripke can justify his acceptance of a certain essentialist assumption that is needed to carry out the proof of the necessity of identity without being forced to appeal to (as Della Rocca claims that he is) the necessity of identity.

Finally, in the third chapter, I assess in detail a challenge to a consequence of the orthodox view —namely, the claim that identity is indeterminate in certain circumstances. Since Evans (1978) established in a famous argument

that all alleged cases of indeterminate identity are really cases of numerical distinctness, my strategy to vindicate the orthodox view in this case consists in defending Evans' argument from a series of criticisms and objections that have been raised against it. Thus, after presenting Evans' argument in the most charitable way, I review a number of objections that have been addressed to it —in particular, some traditional objections put forward by Parsons and Woodruff (1995) and Van Inwagen (1990) and other more recent objections raised Williams (2008) and Barnes (2009)— and I show that all of these objections ultimately founder in virtue of a variety of reasons.

Chapter 2

Rejecting relative identity

2.1 Introduction

Let me start by rehearsing a well-known fact: relativistic motifs and themes have become rather common not only in many social sciences but also in many areas of philosophy. For instance, several authors (e.g., Harman 1977) hold that there are no absolute facts regarding what is morally correct and that, consequently, moral facts only hold relative to different conceptual frameworks. Others (e.g., Feyerabend 2010) maintain that perception cannot be truly objective and impartial because it is always relative to some particular theory. And there are even some (e.g., Putnam 2005) who argue that logical primitives such as the notion of object or that of existence are not absolute but relative to the rejection or acceptance of certain mereological theses.[1]

[1]For a recent critical discussion of some popular relativistic theses (in particular, the relativity of epistemic justification), see Boghossian (2006).

Considering the current prevalence of these relativistic theses, it is not surprising that some philosophers have been led to hold that the identity relation is not absolute but relative in some respects. In particular, Geach (1967) famously argued that there is no unitary and absolute identity relation, but rather a plurality of identity relations that are relative to different sortal concepts. Geach's relative identity theory (henceforth, RIT) has generated an extensive discussion in the philosophical literature, with some philosophers endorsing and defending RIT[2] and others —in fact, the vast majority— arguing that, in addition to being poorly supported, the theory is extremely problematic and has to be consequently rejected.[3]

Now, why has RIT generated such an extensive discussion? What is precisely at stake in the debate over the acceptance or rejection of RIT? Let me touch briefly on some reasons that explain why RIT has generated such controversy. One of the reasons that explain why RIT has generated a heated debate is that, if accepted, it forces us to revise one of the cornerstones of logic —namely, identity theory. In particular, adopting RIT forces us to replace the formal principles that govern classical identity (namely, reflexivity, transitivity and symmetry) by relativized principles that only hold for particular sorts of objects.[4] Accepting these sort-relativized principles would amount to give up the idea that logic is a topic-neutral discipline. To be more specific, for RIT advocates, logical principles such as the reflexivity or

[2]In addition to Geach, some other prominent advocates of RIT include Griffin (1973) and Cartwright (1987).

[3]Some of the most notorious critics of RIT are Wiggins (1967, 2001), Perry (1970, 1978), Dummett (1981) and Feldman (1969).

[4]An exposition of the formal schematic principles that govern relative identity relations can be found in Griffin (1976).

the transitivity of identity do no longer hold true regardless of the objects considered, but are specific to different sort of entities.

Another reason that explains why RIT has been fiercely debated is that it has straightforward implications for the debate on the scope of quantification. Considering that, as Geach (1967: 6) puts it, 'quantification theory and identity theory [are] like two cogwheels that mesh together and move together', accepting RIT forces us to accept that, if identity is not absolute but relative to distinct sorts of entities, there is no all-encompassing domain over which our quantifiers range but rather a series of restricted domains of quantification (e.g., sets, swans, stars, etc.), each of which corresponds to a particular identity relation.

Now, as we can appreciate, several important theses (e.g., the topic-neutral character of logic and the existence of an all-encompassing domain of quantification) hinge on the debate of whether identity is relative or not. In light of this, the issue of whether RIT is true or not is of crucial importance and, accordingly, I intend to evaluate it in detail in the present chapter.

At this stage, some might raise doubts concerning the relevance of going over RIT since doing so appears to be a clear case of flogging a dead horse. Indeed, after Geach initially presented RIT, many authors have critically examined the theory and have concluded that his main arguments in favor of it are flawed. Now, though I agree with the general concensus that Geach's arguments fail to support RIT, I believe that there are stronger arguments that may be deployed in favor of the theory, which have not been addressed yet. Accordingly, my main goal in this chapter is to review critically the case for RIT in light of these arguments and show that, even if we rely on

them (and not on the traditional ones which have been widely discussed), RIT remains quite implausible.

Here is roadmap of the chapter. In §2.2, I present the core of RIT as it is was originally outlined by Geach, paying attention to distinguish all the different theses that Geach puts forward and defends as part of RIT and to highlight the different relations between them. Having done this, I turn in §2.3 to present the main arguments that Geach provides in favor of the central theses of RIT and I go over briefly some of the standard responses that these arguments have elicited in §2.4. Subsequently, I consider in §2.5 some novel arguments that can be used to support one of the central theses of RIT and I provide responses to them Finally, in §2.6, I offer a brief conclusion.

2.2 RIT: the main theses

Though Geach often speaks of RIT as if it involved a single thesis, a close reading of his works reveals that RIT really consists of several interconnected theses. In order to have a clear perspective of RIT, it is important to distinguish clearly these theses and to show what relations hold between them. Following Noonan (1997), one may distinguish in RIT three main theses: (i) the non-existence of classical identity, (ii) the sortal relativity of identity and (iii) the derelativization thesis.[5] In the following subsections, I present the gist of these three theses.

[5]Noonan mentions three other minor theses that Geach endorses as part of RIT: the counting thesis, the irreducibility of restricted quantification and the distinction between a name 'for' an A and a name 'of' an A. I will present and discuss these minor theses further down in §2.3 when I present the arguments given in favor of the three main theses discussed in this section.

2.2.1 The non-existence of classical identity

As Geach (1967: 3) remarks, the notion of classical identity appears to be presupposed in the branch of logic called identity theory. Identity theory is traditionally deemed to be concerned with the identity relation, which is taken to be the binary relation that every thing bears to itself and to nothing else. In light of this, identity is an equivalence relation (i.e., it is reflexive, transitive and symmetric) that abides by the Indiscernibility of the Identical. Now, considering that identity has all these features, some logicians maintain that it can be uniquely characterized in terms of the following schema (where '=' is a two-place predicate belonging to a language L with a fixed stock of constants, predicates, quantifiers and truth-functional connectives):

(ID): $\vdash Fa \leftrightarrow (\exists x)(Fx\& \ x{=}a)$

Indeed, any two place-predicate denoting a binary relation that satisfies ID denotes an equivalence relation in L that is governed by the Indiscernibility of the Identical. But a serious difficulty looms here according to Geach. Even if the fact that two distinct two-place predicates satisfy separately ID is enough to guarantee that these two predicates denote the same relation, there is no guarantee for Geach that a two-place predicate R that satisfies ID in a language L really expresses identity rather than some other relation. In particular, Geach (1967: 5) maintains that there is no guarantee that R expresses, rather than strict identity, mere indiscernibility by the predicates that make up the expressive resources of L.

To grasp more clearly Geach's point, it is important to keep in mind the historical background in which it was presented. This background is domi-

nated by one of the key projects undertaken by Quine in his work —namely, the 'limning [of] the true and ultimate structure of reality' (1960: 221)— and by the specific way in which Quine carried this project. Indeed, even though philosophers such as Spinoza or Leibniz also undertook this project, Quine differs radically from them in virtue of the methodology that he adopts: whereas the inquiries of his predecessors into the structure of reality were driven by a series of *a priori* reflections, Quine (1981: 21) maintains that 'it is within science itself, and not in some prior philosophy, that reality is to be identified and described'. To be more specific, Quine argues that our best scientific theories provide us, at a given time, with the best picture of the true and ultimate structure of reality. This may leave some yearning for more in the sense that scientific theories are always open to revision and that, accordingly, we lack an absolute standpoint to assess their correctness. In response to this, Quine (1960: 3) replies that our belief in the existence of this absolute standpoint stems from a mistake that consists in 'seeking an implicit sub-basement of conceptualization, or of language'. Once we acknowledge that there is no such absolute standpoint and that philosophers and scientists are on the same boat, Quine's emphasis on the role of science as the primary tool we have investigate reality becomes a far more plausible thesis.

But even if we grant that our best scientific theories provide us with the best picture of reality, some problems remain. In particular, although our best scientific theories involve theoretical terms such as 'electron', 'gene' or 'mass', they also involve others terms such as adjectives (e.g., 'large', 'red' and 'tall'), adverbs (e.g., 'near', 'almost' and 'always') and grammatical con-

junctions (e.g., 'and', 'or' and 'so'). If we consider exclusively non-theoretical terms such as the aforementioned ones,[6] it is clear that they are subject to phenomena that Quine labels 'vagaries of reference': some of them exhibit vagueness, others ambiguity and there are also cases of referential failure. To avoid (or, at least, limit) the impact of these linguistic phenomena on the realization of the project, Quine suggests regimenting our scientific theories —i.e., paraphrasing them in logical notation.

In addition to the fact that it enables us to eliminate ambiguities and other problematic features that are present in natural languages, Quine (1960: 158) argues that there is another strong motivation to offer a regimentation of scientific theories: it allows us to simplify the theories in various ways. In particular, Quine mentions that regimentation allows us to minimize the stock of primitive functions or constructions within the theories. It is in connection with this benefit of regimentation that Geach's remarks on the lack of any guarantee that that R expresses identity must be considered.

Indeed, whereas many distinguished logicians and philosophers take the notion of identity as primitive,[7] others such as Quine himself maintain that one of the best reasons we have to consider identity as a part of logic (rather than as an extra-logical predicate belonging to mathematics) is that we can offer a reduction of identity in terms of certain expressive resources that

[6]Several authors (e.g. Hanson 1958) have persuasively argued that there is no clear distinction between theoretical and non-theoretical terms. Though I am sympathetic to this view, I will assume here that there is a distinction just for the purposes of fleshing out more clearly Quine's views

[7]Frege is a paradigmatic case since he (1894: 320) argues in his review of Husserl's *Philosophy of Arithmetic* that, given that any definition expresses an identity of content between the *definiens* and the *definiendum*, identity itself cannot be defined. Similar views are defended by Savellos (1990), McGinn (2000) and Hawthorne (2006).

belong unquestionably to logic (such as quantifiers, variables and truth-functional connectives). In order to appreciate how this can be done, let me rehearse an example presented by Quine himself:

> Consider a standard language whose lexicon of predicates consists of a one-place predicate 'A', two two-place predicates 'B' and 'C' and a three-place predicate 'D'. We then define '$x{=}y$' as short for:
>
> $$\mathrm{A}x{\equiv}\mathrm{A}y \cdot \forall z[\mathrm{B}zx{\equiv}\mathrm{B}zy \cdot \mathrm{B}xz{\equiv}\mathrm{B}yz \cdot \mathrm{C}zx{\equiv}\mathrm{C}zy \cdot \mathrm{C}xz{\equiv}\mathrm{C}yz \cdot \forall z'(\mathrm{D}zz'x \equiv \mathrm{D}zz'y \cdot \mathrm{D}zxz'{\equiv}\mathrm{D}zyz' \cdot \mathrm{D}xzz'{\equiv}\mathrm{D}yzz')]$$
>
> Note the plan: the exhaustion of combinations. What '$x{=}y$' tells us, according to this defintion, is that the objects x and y are indistinguishable by the four predicates; that they are indistinguishable from each other even in their relations to any other objects z and z', insofar as these relations are expressed in simple sentences. (1986: 63)

As we can appreciate, Quine's characterization of identity involves defining, within a certain language L, a relation of equivalence such that any two objects that satisfy it are indiscernible by the expressive resources of L. Quine argues that this relation of equivalence may be treated as identity in L if we are willing to reinterpret the expressive resources of L accordingly. To illustrate this, Quine considers the following example: he supposes that there is a certain language that makes reference only to persons and he also supposes that, within this language, whatever is true of any person is also

true of any other person that has the same income as the first one. If this is the case, we may then, as Quine (1963: 71) suggests, treat the two-place predicate 'x has the same income as y' as identity by reinterpreting 'the references of the original objects [...] for purposes of the discourse as referring to other and fewer objects, in such a way that indistinguishable originals [i.e., persons with the same income] give way to the same new object [i.e., an income group].'

A crucial point that Quine (1986: 63) himself emphasizes is that the method he proposes to define identity within a certain language L only works provided the lexicon of L is finite: there is no guarantee that, if the lexicon is infinite, we can effectively characterize identity using the method. In fact, Quine makes a further concession: he acknowledges that it may happen that the objects intended as the values of the variables x and y in L are distinct but not completely distinguishable from each other by the predicates of L. If this occurs, identity cannot be defined in terms of a formula that expresses indiscernibility between the values of x and y in L, though this shortcoming is unnoticeable within L. This is precisely the limitation that Geach exploits to argue that we have no guarantee that any two-place predicate defined in accordance to Quine's method in a certain language L expresses genuine identity rather than mere indiscernibility in L.

Now, if we have no guarantee that any two-place predicate defined using Quine's method expresses genuine identity rather than some other relation in a language L, there is a good reason to maintain that there is no classical identity, particularly in light of the view that Geach defends elsewhere regarding concepts. Indeed, since Geach (1957) maintains that concepts (such

as the one expressed by a two-place predicate that satisfies ID in a certain language L) are mental entities the primary purpose of which is to be 'exercised' in judgments in such a way that they determine a particular content, it is clear that, if a certain concept fails to be properly 'exercised' in a judgment (i.e., if the expression that allegedly denotes it fails to correspond to a determinate content),[8] there is a good motivation to maintain that the concept in question has no determinate content. Hence, if there is no expression that can effectively refer to the concept of classical identity without restrictions, there is a good motivation according to Geach to maintain that there is nothing that corresponds to the concept of classical identity (or, in other terms, that there is no classical identity).

To recap, the first thesis that Geach endorses as part of RIT can be formulated in the following way:

> (RIT_1) Given that there can be no expression designating classical identity in any possible language, the concept of classical identity has no determinate content and must be treated as having no instances.

Besides RIT_1, there are two other crucial theses that Geach adopts in his presentation and defense of RIT. I turn to these in the next subsections.

2.2.2 The sortal relativity of identity

Granting that there is indeed no strict, absolute identity, the question that arises immediately is the following: "What does the two-place predicate

[8] Geach (1969: 556) emphasizes very clearly this about the concept of identity when he writes: "I deny that any *one* relation [...] is *the* relation expressed by 'is identical with'."

'x is identical with y' stand for?" To this question, Geach suggests two possible answers: either it is a vague expression that fails to denote anything determinately or it is an expression that is shorthand for another predicate, which has the form 'x is the same A as y'. Thus, according to Geach, if we are to make sense of sentences involving the expression 'x is identical with y', we must construe them as expressing, not absolute identity, but rather a relation of relative identity which is defined by reference to a certain general term A that is supplied by the context.

What evidence does Geach rely on to argue that 'x is identical to y' is an expression that stands for an identity predicate that is relativized to a certain general term? After highlighting the close connection between the notions of identity and cardinality (which is manifested in the fact that statements that contain assertions of identity such as 'a is *identical* with b' entail statements that contain assertions of cardinality such as 'a and b are just *one*'), Geach appeals to Frege's view on cardinality to defend the sortal relativity of identity:

> Frege emphasized that 'x is one' is an incomplete way of saying 'x is one A, a single A' or else has no clear sense; since the connection of the concepts *one* and *identity* comes out just as much in the German '*ein und dasselbe*' as in the English 'one and the same', it has always surprised me that Frege did not similarly maintain the parallel doctrine of relativized identity, which I have just briefly stated. (1967: 3)

For Frege, the fact that the expression 'x is one' is incomplete stems from his view on statements of number. Indeed, given that a statement of

number expresses according to him (1884: 59) an assertion about a concept, it is clear that a statement such as 'Washington and Houston are two' is not well-formed for Frege because it fails to specify which concept the number two falls under. It is because of this view of numerical attributions that Frege (1884: 40) maintains that the expression 'one man' cannot be treated as the expression 'wise man': we cannot split the expression 'one man' to obtain two meaningful predicates as we can split the expression 'wise man'. Now, using the aforementioned connection between the notions of identity and cardinality and making again reference to Frege's view, Geach provides elsewhere another statement of the thesis of sortal relativity of identity that illustrates a further feature of it:

> Frege has clearly explained that the predication of 'one endowed with wisdom' ('ein Weiser') does not split up into predications of 'one' and 'endowed with wisdom' ('weise'). It is surprising that Frege should on the contrary have constantly assumed that 'x is the same A as y' does split into 'x is an A (and y is an A) and x is the same as (*ist dasselbe wie*, *ist gleich*) y'. (1962: 151-150)

As we can appreciate, Geach maintains that, in addition to the fact that the expression 'x is identical to y' really stands for a relative identity predicate of the form 'x is the same A as y' (where A is a general term provided by the context), the relative predicate in question is primitive: it cannot be decomposed or broken down into more basic components. In particular, it does not admit being analyzed into a two-place predicate expressing absolute identity and a series of predications involving the general term A.

A crucial element in the thesis of the sortal relativity of identity that Geach endorses is the general term A that is used to determine the relative identity relation denoted by an instance of the two-place predicate 'x' is identical with y'. This general term serves various purposes. First and foremost, it makes identity statements meaningful by enabling us to specify which relative identity relation they refer to; without it, there is simply no telling which identity relation is picked out by the statement and, consequently, whether we have asserted in uttering the statement something true or false. Indeed, considering that Geach (1962: 39) holds that "it makes no sense to judge whether x and y are 'the same', or whether x remains 'the same' unless we add or understand some general term —the same A", it is then clear that, unless the identity statement in question involves (either implicitly or explicitly) a general term, it cannot be assigned a truth-value.

Besides making meaningful the identity statements to which it is associated, the general term also has other functions. As we previously mentioned, there is a close connection between the notions of identity and cardinality in the sense that accepting certain identity statements (e.g., 'Mark Twain is *identical* with Samuel Clemens') entails accepting certain cardinality statements (e.g., 'Mark Twain and Samuel Clemens are *one*'). Now, it is clear that a meaningful cardinality statement has to provide us with a precise count of the items it refers to. But this, as Frege remarked, can only be done if the cardinality statement in question involves a certain general term:

> If I place a pile of cards in his hands with the words: Find the number of these, this does not tell him whether I wish to know the number of cards, or of complete packs of cards, or even say

> of points in the game of skat. To have given him the pile in his hands is not yet to have given him the object he is to investigate; I must add some further word —cards, or packs, or points. (1884: 28-29)

For Geach, a second role of the general term attached to a relative identity predicate is to provide us with a criterion to count things. It is because of this reason that the general terms that are used to specify relative identity predicates are deemed to express *sortal* concepts.[9] But enabling us to provide a count of things is not the only function of the general term associated to a relative identity predicate. In addition to this, Geach (1962: 39) maintains that the general term also has another role —namely, to provide a *criterion of identity*.

The notion of a criterion of identity has played a crucial role in many philosophical discussions throughout history. Unfortunately, there is also a great deal of controversy surrounding this notion in the sense that it has often been taken to be different things and to perform different functions.[10] In particular, there are at least two ways to understand the notion of a criterion of identity that have been extremely influential. The first one, which surfaces in Locke's writings, has a key *metaphysical* import: a criterion of identity for an object o belonging to a certain kind F provides a measure of what is

[9]Several authors (e.g., Strawson 1959 and Wallace 1965) characterize a sortal concept as a concept that provides a criterion for counting items of a certain kind. This characterization of sortal concepts, however, is not the only one used in the philosophical literature. See footnote 12 below for further discussion and references.

[10]A thorough assessment of the notion of criterion of identity can be found in Lowe (1989b). Williamson (1986) and Zimmerman (1998) also provide valuable discussions of the different forms and functions that criteria of identity have.

required to be o.[11] In particular, Locke (1995: 243) relies on this sense of the notion of criterion of identity to argue that what makes an oak be an oak is different from what makes a mass of plant matter be a mass of plant matter. The second way of understanding the notion of a criterion of identity has, on the contrary, an *epistemological* import: a criterion of identity allows us to recognize that a certain object that is picked by a certain description is identical to an object that falls under another description. This second way of understanding the notion of criterion of identity is clearly present in Frege's works, as the following passage illustrates:

> If we are to use the symbol a to signify an object, we must have a criterion for deciding in all cases whether b is the same as a, even if it is not always in our power to apply this criterion. (1884: 73)

As we can appreciate, we have two different ways of understanding the notion of criterion of identity. Now, since Geach maintains that a general term that specifies a relative identity relation provides us with a criterion of identity (and is then, accordingly, a sortal),[12] the question that naturally arises is this: 'What role does the criterion of identity conveyed by the general

[11] In light of this, one could say that a criterion of identity for an object o thus understood provides us with the (nominal) essence of o. Cf. Yablo (1987: 297).

[12] In addition to the characterization of a sortal concept as a concept that allows us to provide a *count* of items of certain kind, a sortal concept is also often characterized as a term that conveys a *criterion of identity* for entities of certain kind. There are also further characterizations of the notion of sortal concept in mereological terms and in terms of essences, as Feldman (1973) and Grandy (2007) remark. For the sake of simplicity, I will endorse here, following Lowe (1989a: 2), a characterization of sortal concept that only involves the two aformentioned features (which Geach accepts): a sortal concept provides both a criterion for counting items of a certain kind and a criterion of identity for items of a certain kind. This characterization rules out *mass terms* (water, sand, gold, etc.) as expressing sortal concepts since mass terms convey a criterion of identity but not a criterion for counting.

term associated to a relative identity predicate have according to Geach?' Regarding this question, Geach does not take a very clear stance. Indeed, he seems to adopt in certain passages the metaphysical interpretation of the notion of criterion of identity to the extent that he maintains that an object a may either be identical to an object b or not depending on which criterion of identity we focus on:

> On my own view of identity, I could not not object to different A's being one and the same B, and thus different intentional objects could be one and the same man, as different official personages may be one and the same man. (1962: 157)

However, there are other passages where Geach appears to embrace clearly the epistemological construal of the notion of criterion of identity. Given that he (1973: 289) argues elsewhere that 'it is as nonsensical to speak of identification apart form identifying some kind of thing as to speak of counting apart from counting some kind of thing', the role that he attributes to a criterion of identity consists in prodiving us a standard to recognize objects as either being identical to each other or distinct from each other.

Thus, despite our best efforts, it is difficult to provide a clear formulation of the thesis of sortal relativity of identity because Geach himself wavers between different views concerning the role of the criterion of identity associated to a certain relative identity predicate. In light of this, it is convenient to formulate two alternate versions of the thesis of sortal relativity as possible candidates for being included in RIT. If we consider the criterion of identity conveyed by a sortal associated to a certain relative identity predicate to

have a metaphysical role, the thesis of sortal relativity of identity may be formulated in the following way:

> (RIT_{2m}) An identity statement 'a=b' really expresses a relative identity statement of the form 'a is the same F as b' that does not admit being analyzed in more basic terms and in which 'F' is a general term such that (i) a and b both fall under the concept it expresses, (ii) the concept expressed by 'F' provides a criterion to count entities of some kind K to which both a and b belong and (iii) the concept expressed by 'F' provides us with a criterion of identity that specifies what it takes for a to be identical to b.

Now, considering that Geach admits that one individual can belong to several different kinds (for instance, a certain individual may be both a mass of plant matter and an oak) and that different criteria of identity may correspond to these kinds, it it clear then that, if criteria of identity are primarily understood as having a metaphysical role, a consequence of the thesis of sortal relativity of identity is a certain picture of reality that Dummett describes in the following terms:

> [...] Geach means us to picture that over which the variables range as an amorphous lump or reality, in itself not articulated into distinct objects. Such an articulation may be accomplished in any one of many different ways: we slice up reality into distinct individual objects by selecting a particular criterion of identity. (1981: 563)

There are a number of difficulties associated with this picture, but I will address them later. My main concern in this subsection is to present as clearly as possible the main theses that make up RIT —or, at least, the most plausible ways to interpret the main theses that make up RIT. Now, if we construe the notion of criterion of identity as having primarily a metaphysical role, the best formulation of the thesis of sortal relativity of identity is RIT_{2m}. But, if we construe the notion as having primarily an epistemological role, a more appropriate formulation of the thesis is the following:

> (RIT_{2e}) An identity statement 'a=b' really expresses a relative identity statement of the form 'a is the same F as b' that does not admit being analyzed in more basic terms and in which 'F' is a general term such that (i) a and b both fall under the concept it expresses, (ii) the concept expressed by 'F' provides a criterion to count entities of some specific kind to which both a and b belong and (iii) the concept expressed by 'F' is associated with a certain criterion of identity that provides us with a way to identify a with b.

Having formulated two possible ways to construe the thesis of sortal relativity of identity, we may consider presently the issue of what relation they have with respect to RIT_1. Noonan (1997: 637) has argued that the non-existence of classical identity fails to entail either RIT_{2m} or RIT_{2e} and, conversely, that no interpretation of the thesis of sortal relativity of identity entails RIT_1. Though I agree with him that the theses are logically independent (i.e., one can consistently accept RIT_1 and reject both interpretations of the thesis of sortal relativity or accept any of the interpretations of the

thesis of sortal relativity and reject RIT_1),[13] there is nevertheless a very close link between them. As I see things, the link is the following: if RIT_1 is true, we have strong motivations to maintain that one of the interpretations of the thesis of sortal relativity of identity is true as well. Indeed, if it is the case that no expression in any language whatsoever can refer to classical identity but we do not wish to dismiss '$x{=}y$' as a meaningless string of symbols, the only alternative that seems to be available to us is to construe the expression as being an abbreviation of a relative identity predicate. It is important to notice that accepting RIT_1 does not, by itself, provide a stronger motivation to embrace RIT_{2m} over RIT_{2e} (or viceversa): RIT_1 provides equal support to both RIT_{2m} and RIT_{2e}. If a choice needs to be done between the two formulations of the thesis of sortal relativity of identity, it must be done on the basis of considerations independent from their relation to RIT_1.

2.2.3 The derelativization thesis

We have seen in §2.2.2 that there are two possible interpretations of the thesis of sortal relativity of identity. Regardless of the interpretation we consider, an important fact to bear in mind is, as Noonan points out, that the significance of the thesis of sortal relativity presupposes a distinction between sortal terms (i.e., terms expressing concepts that convey both a criterion to count and a criterion of identity) and non-sortal terms. Now, to make

[13] At this stage, someone might be tempted to ask the following question: is there a consistent position that admits both interpretations of the thesis of sortal relativity of identity? To the best of my knowledge, Geach's opponents and his supporters are silent on this issue. However, considering that Geach (1973: 288) accepts the distinction between the two ways to understand the notion of criterion of identity as a genuine distinction (he explicitly states that 'it is not illusory'), we have reasons to believe that RIT_{2m} and RIT_{2e} should not be endorsed at once, under pain of conflating them.

sense of this distinction, which reflects a distinction between *substantival* and *adjectival* terms that Aquinas draws according to Geach (1962: 39),[14] we need an account of how we can grasp sortal terms and differentiate them from non-sortal terms. Geach's suggestion is the following: since relative identity predicates such as 'x is the same F as y' are primitive, one may account for our understanding of sortals terms by arguing that we obtain them by a process of *derelativization*, which Geach characterizes using the following example:

> [...] we can offer 'x is the same A as *something*' to analyze away the predicative use 'x is an A' of a substantival general term. For our using the expression 'x is the same A as' requires 'A' to be construed as such a term, and in that case a thing is an A if and only if it is the same A as something or other [...]. (1962: 191)

As we can appreciate, the aforementioned process of derelativization provides an account of how we can understand sortal terms and distinguish them from non-sortal terms: we first consider relative identity statements of the form 'a is the same F as something' (where F is a general term) and, having done that, we determine if the statement 'a is an F' (which is obtained by derelativizing the identity statement) makes sense. If that is the case, the term F is a sortal term. Consequently, a sufficient condition for a general

[14] The distinction between substantival and adjectival terms does not mirror exactly that between sortal terms and non-sortal terms because Geach admits that mass terms such as 'gold' are substantival on the basis that they admit plural forms and that they convey a criterion of identity even if they do not provide a criterion to count. Dummett (1981: 201-202) has argued that treating the expression 'substantival term' as synonym for 'sortal term' threatens to obliterate the distinction between mass terms and sortal terms, so I will avoid the expression 'substantival term' whenever possible.

term F being a sortal according to Geach is that the predicate 'x is an F' can be obtained from the relative identity predicate 'x is the same F as something' by derelativization.[15]

The thesis that sortal terms arise from a derelativization process from relative identity predicates is partially based on certain views of Quine, as Noonan remarks. Indeed, after he establishes a distinction between terms that are true of objects considered in isolation (such as the predicate 'x is a set') and terms that are true of objects considered in relation some other objects (such as the predicate 'x is a brother of y'), Quine maintains that 'relative terms', which belong to the second category, can in certain circumstances be used as if they were true of true of objects in isolation through a process of derelativization:

> Commonly, the key word of a relative term is used also *derelativized*, as an absolute term to this effect: it is true of anything x if and only if the relative term is true of x with respect to at least one thing. Thus, anyone is a brother if and only if there is someone of whom he is a brother. (Quine 1960: 106)

To grasp in more detail the notion of derelativization thesis that Quine refers to in the abovementioned passage, it is important to keep in mind the importance of regimentation for the achievement of Quine's project. As I

[15]Geach actually suggests much a broader thesis: substantival terms (and not merely sortal terms) are obtained by derelativization from predicates expressing relative equivalence relations. Thus, the substantival term 'x has a temperature T', which conveys a criterion of identity but not a criterion to count, is obtained for Geach by derelativizing the predicate 'x has the same temperature T as y'. Considering the potential problems raised by the notion of a substantival term (see footnote 14), I will focus here exclusively on sortal terms.

mentioned in §1.2.1, Quine argued that we should regiment our best scientific theories by drawing attention to the clarity and simplicity that they gain when they are paraphrased in a purely logical notation that involves predicates, quantifiers and variables. In fact, the quest for simplicity eventually led him to maintain that variables and quantifiers can be safely dispensed with: all reference to particulars in our scientific theories can be preserved by means of a universal algebra that involves exclusively a series of n-ary predicates and six operators.[16] Among these operators, there is one that occupies a preeminent role —namely, the *derelativization* operator (which Quine symbolizes using the expression 'Der'). To illustrate how the operator works, Quine asks us to consider initially a two-place predicate 'Bxy' (which stands for 'x bites y') and he then explains how it can be transformed into a one-place predicate in the following passage:

> [...] we need an operator that can be applied to a two place-predicate or transtive verb 'B', 'bites', to produce a one place-predicate 'x bites something'. Let us call this operator that of *derelativization* and write it 'Der'. Thus '$Der\ B$' is the one-place-predicate or intransitive verb of biting, or biting something, and the predication '$(Der\ B)x$' means that x bites something. (1960: 344)

As we can appreciate, the operator of derelativization provides us with the basic mechanism to achieve a better regimentation of our theories by

[16] I will not rehearse here Quine's proposal to dispense with variables in its entirety since this would take up too much space. I will only present what is required to understand how Geach introduces and justifies the derelativization thesis.

eliminating variables. Now, since variables are traditionally used (as Quine readily acknowledges) to mark positions, it is crucial for him to be able 'to coax variables into the right position', which boils down to be able to distinguish the relation expressed by the predicate 'Bxy' from the relations expressed by the predicates 'Byx' and 'Bxx'. To do this, Quine introduces two other operators, Inversion (which he writes 'Inv') and Reflection (which he writes 'Ref'), that he characterizes in the following way: '$(Inv\ B)xy$' stands for 'Byx' and '$(Ref\ B)x$' stands for 'Bxx'.

Using Quine's notion of derelativization, Geach argues that relative identity predicates can be used in a derelativized way. Of course, the derelativized use only makes sense according to him because the relative identity predicates are taken to be primitive expressions that admit no analysis. Now, given that the significance of the derelativization thesis depends on the primitiveness of relative identity predicates (which is part of the thesis of sortal relativity of identity), an interesting question arises: assuming that the predicate 'x is the same F as y' is indeed primitive (i.e., that it is not analyzable as 'x is an F, y is an F and x is the same as y'), what is the function of the expression 'the same' in the relative identity predicate? Geach provides an answer to this question that is in agreement with the derelativization thesis:

> We shall treat 'the same' in 'is the same A as' not as a syntactically separable part but as an index showing we have here a word for a certain sort of relation: just as 'of' in 'is a brother of' does not signify a relation by itself (as if the phrase were 'is a brother, who belongs to'), but serves to show that the whole, 'is a brother of', stands for a relation. (1973: 291)

In a nutshell, what Geach suggests then is that, if we take relative identity predicates of the form 'x is the same F as y' as primitive expressions and if we use Quine's proposal (or something similar) to symbolize them, we can appreciate why sortal predicates such as 'x is an F' can be analyzed in terms of certain corresponding relative identity predicates. Indeed, if we grant that the sortal predicate 'x is an F' is equivalent (as Geach argues) to the one-place predicate 'x is the same F as something' and this one-place predicate in turn results from the derelativization of the relative identity predicate 'x is the same F as y', it is clear that sortals may then be understood in terms of certain relative identity predicates.

After showing how Geach's derelativization thesis involves some key notions that Quine developed in his attempts to produce the best possible regimentation of scientific theories, we can now provide a clear formulation of the thesis and show its connections with the other two theses that make up the core of RIT. Here is a possible formulation of the thesis:

> (RIT_3) Any sortal predicate of the form 'x is an F' can be obtained from a relative identity predicate of the form 'x is the same F as y' by a process of derelativization in which the relative identity predicate in question is transformed into a one-place predicate of the form 'x is the same F as something', which is equivalent to the sortal predicate 'x is an F'.

Now, since the significance of the thesis of sortal relativity of identity presupposes a distinction between sortal and non-sortal terms and this requires in turn an account of how sortal terms are understood, it seems that

the truth of RIT_3 is a necessary condition for RIT_2 in any of its two formulations. Moreover, there also seems to be an intimate connection between RIT_1 and RIT_3, which can be expressed in the following terms: since the fact that no expression stands for classical identity suggests that it does not make sense to analyze the expression 'x is the same F as y' as 'x is an F, y is an F and x is the same as y', the truth of RIT_1 sanctions the possibility of reversing the traditional direction of analysis and treat sortals terms as being obtained from primitive relative identity predicates, which is the gist of RIT_3.

2.3 Geach's arguments for RIT

After presenting in the previous section the main theses that constitute RIT and outlining the connections existing between them, I turn now in this section to rehearse and evalauate the main arguments that have been provided to support them. Following the order of exposition that I adopted in the previous section, I will consider successively the arguments that Geach presents for RIT_1, RIT_2 and RIT_3 while providing an overview of the relations between these theses and three other minor theses that Geach endorses as well.

2.3.1 Geach's arguments for RIT_1

In order to establish RIT_1, Geach deploys the following argumentative strategy. He asks us to consider a two-place predicate I that belongs to a language L and he stipulates that this predicate satisfies the schema ID for all

predicates of L. Having done that, he observes that, though this predicate (which he calls an 'I-predicate') may express classical identity in L, there is nothing that guarantees this fact. Using Quine's admission that an I-predicate defined in a language L may really express indiscernibility between distinct objects rather than identity and that this limitation is unnoticeable if we consider exclusively the expressive ressources of L,[17] Geach argues that, in order to guarantee effectively that the I-predicate expresses identity, we need somehow to be able to characterize the I-predicate in a language-transcedent fashion but any attempt to do this creates a serious problem:

> 'For real identity, though', we may wish to say, 'we need not bring in the ideology of a definite theory T. For real identity, *whatever* is true of something identical with a is true of a and conversely, regardless of which theory this may be expressed in; and a two-place predicable signifying real identity must be an I-predicable no matter what other predicables occur with it in the theory.' But if we wish to talk this way, we shall soon fall into contradictions; such unrestrained language about 'whaterver is true of a', not made relative to the definite ideology of a theory T, will land us in such notorious paradoxes as Grelling's and Richard's. If, however, we restrict ourselves to the ideology of a theory of T, then, as I said, an I-predicable need not express identity, but only indiscernibility within the ideology of T. (1967: 5)

[17] Geach talks of I-predicates being defined with respect to a certain *theory* rather than with respect to a certain *language*. The rationale that he (1973: 298) gives for this is that whereas 'a language normally contains all the negations of its sentences, a theory will not contain the negation of all of its theses.' Following Geach, I consider here I-predicates being defined with respect to theories, although this creates problems (as I show below).

As the previous passage clearly shows, Geach argues for RIT_1 by setting up a dilemma that advocates of classical identity face. The dilemma is presented in the following way: after introducing the notion of an I-predicate in a certain theory T as a two-place predicate that satisfies the schema ID in T and reminding us that several logicians (in particular, Geach mentions Hao Wang) maintain that ID is enough to characterize identity since it may serve as a single axiom schema for identity theory, Geach presents the two horns of the dilemma: (i) it might be the case that the I-predicate in question does not really express identity in T but rather mere indiscernibility vis-à-vis the expressive resources of T (a fact that cannot be appreciated within the framework of T) because the I-predicate is defined relative to T and (ii) any attempt to characterize an I-predicate in a theory-independent way will force us to embrace a view in which problems similar to certain semantic paradoxes arise. And, since neither horn of the dilemma is acceptable for an advocate of classical identity, Geach then concludes that no expression (and, more particularly, no I-predicate) can refer to classical identity. Here is a more detailed formulation of the argument:

> (1) An I-predicate in a theory T is an expression that satisfies the schema ID for all the predicates of T. (Definition)
>
> (2) Since ID can be used as the basic axiom of identity theory, this suggests that, if identity can be expressed in T, it is expressed by a certain I-predicate belonging to T. (Wang's hypothesis)
>
> (3) Any I-predicate is such that either it is defined in relation to a certain specific theory T or it is defined in a theory-independent way. (Tautology)

(4) If an I-predicate is defined in relation to T, we have no guarantee that it refers to identity in L because the expressive resources of T do not allow us to distinguish cases of mere indiscernibility in T from cases of genuine identity in T. (Rejection of the first horn of the dilemma)

(5) If an I-predicate is defined in a theory-independent way, we end up endorsing a view that is undermined by problems akin to those raised by semantic paradoxes such as those presented by Richard and Grelling. (Rejection of the second horn of the dilemma)

(6) Consequently, we may conclude that there is no expression (and, in particular, no I-predicate) that denotes identity, which in turn suggests that there is no classical identity.

As this reconstruction of the argument enables us to appreciate, Geach's case for the non-existence of classical identity crucially depends on showing that the two possible alternatives he considers for characterizing I-predicates (i.e., either in relation to certain specific theories or in a theory-independent way) are problematic. Now, in the case of the first alternative, Geach acknowledges that Quine may attempt to resist this argument by adopting a particular strategy —namely, the one I described in §2.2.1. Given that the main problem that arises in the case of the first alternative is the incapacity to distinguish cases of mere indiscernibility in T from cases of genuine identity in T, Quine may appeal to his method of reduction to define cases of indiscernibility as cases of identity.

We have noticed previously that the significance of Quine's method of reduction of identity requires being able to make some changes to the domain that we attribute to our quantifiers. Indeed, since identity and quantification are (as Geach mentions) like cogwheels meshed with each other and moving in unison, treating indiscernible objects in a certain theory (i.e., persons with the same income) as if they were identical clearly involves switching the domain over which our quantfiers range: we no longer consider the orginal individuals (e.g., persons) as values for our variables, but rather certain equivalence classes defined over the original individuals (e.g., income groups) as constituting the new domain over which our quantifiers range.

Geach manifests very serious reservations vis-à-vis the implementation of Quine's method of reduction. These reservations essentially stem from the thesis, which Quine endorses in numerous places, that regimentation enables us to grasp more clearly which entities our best theories are committed to in order to be true.[18] In light of this view, Geach argues that, if Quine were to deploy his method of modifying the ranges of the quantifiers to ensure that a certain I-predicate in a theory T expresses identity in T rather than mere indiscernibility, the reinterpreted sentences of T that involve the I-predicate would have different ontological commitments.

In order to illustrate clearly this point, Geach asks us to consider an universe of discourse that involves exclusively the words written in a book in one of his shelves and an impoverished theory T the expressive resources of which are unable to distinguish between different tokens of certain type-words that occur in the book. He further stipulates that T contains a certain I-predicate

[18]For Quine's views on the role of regimentation as a tool to reveal the ontological commitment of sentences belonging to certain theories, see Quine (1948).

and, having done that, he notices that this I-predicate may be given at least two interpretations. We may construe the I-predicate as expressing mere indiscernibility between the token-words of the book but we may also construe it, if we use Quine's proposal, as expressing identity. Of course, if we choose the second construal, we have to modify the range of the quantifiers of T in accordance with the interpretation assigned to the I-predicate: the variables no longer range over token-words but over type-words. Consequently, in order to be true, sentences of T that include the I-predicate presuppose the existence of these classes. Using this example, Geach argues that the deployment of Quine's reinterpretation method to guarantee that the I-predicate expresses identity rather than mere indiscernibility clashes with some other principles that Quine endorses in the following passage:

> I can show, I think, that this way of interpreting the quantifiers so as to get an absolute identity involves a sin against a highly intuitive methodological program, clearly enunciated by Quine himself [...] The program is that as our knowledge expands we should unhesitatingly exapand our ideology, our stock of predicates, but should be much more wary about altering our ontology, the interpretation of our bound-name variables. (1967: 8)

As we can appreciate, the objection that Geach raises consists in the fact that the reinterpretation of an I-predicate in T to ensure that it expresses identity forces Quine to change the ontological commitments of the sentences of T that contain this I-predicate. In fact, according to Geach, the use of Quine's method is extremely problematic not just because it introduces certain alterations in the ontology required to make the sentences of T

involving the I-predicate true, but because it opens the door to a plethora of new entities, and this clashes with Quine's avowed preference for ontological conservatism. To show more precisely how the application of Quine's method leads to an inelegant proliferation of entities, Geach puts forward the following argument:

> Let us consider a theory T, whose quantifiers are assigned a certain range by Quine's method: if T is a rich theory, we may find within T a number of predicables E^1, E^2, E_3..., each of which is an I-predicable with respect to some corresponding sub-theory of T^1, T^2, T_3..., and none of which has the same application in T as any other. On Quine's view, E^1 will express strict identity in T^1, and E^2 in T^2, and E^3 in T^3..., and the quantifiers of the theories T^1, T^2, T_3,..., must accordingly be reconstrued as having different ranges in each case. We wanted to keep our ontology comparatively fixed while allowing certain changes in our ideology, but now some quite trivial shifts in ideology —the mere omission from some predicables from a theory— will result in quite large additions to our ontology. (1967: 9)

As we can appreciate, the gist of the argument that Geach relies on to criticize Quine's method is that, if we were to employ it to construe each I-predicate E^n that belongs to a certain sub-theory T^n (which is part of a richer theory T) as expressing identity and if we were consequently forced to make corresponding modifications to the interpretations of the remaining predicates of T^n and to the ranges of the quantifiers of T^n, we end up being committed to the entities required to make the sentences of T^n true

in addition to the entities required to make the sentences of T true because T^n is part of T. This argument against the use of Quine's method may be formulated more clearly in the following way:

> (1) Suppose that, from a certain theory T (which involves certain sentences that have specific ontological commitments), we can carve out as parts T^1, T^2, T^3..., and that each one of these subtheories includes a predicate that happens to be an I-predicate for that theory (i.e., it satisfies ID in that theory). (Assumption)
>
> (2) Suppose that we use Quine's method on each one of these subtheories T^1, T^2, T^3... and that, as a result, each I-predicate I^n belonging to a certain theory T^n carved out from T is construed as expressing identity in T^n rather than mere indiscernibility between different objects. (Application of Quine's method)
>
> (3) Even though the application of Quine's method involves modifying both the interpretation of the predicates of T^n and changing the range of the quantifiers of T^n, the truth-conditions of the reinterpreted sentences containing the I-predicate in T^n remain the same as they are in the richer theory T.[19]

[19]Geach defends this premise by considering the case of a theory T that encompasses exclusively the words of a certain volume in his shelves and a theory T^1 that is a subtheory of T. After stipulating that T^1 is unable to distinguish distinct tokens of the same type-word (while T is able) and that T^1 contains both the I-predicate 'Exy' and the predicate 'F' (which stands for "contains two occurrences of 'e' "), Geach (1967: 8) argues that, even if the formula '$(\exists x)(\exists y)(Fx \ \& \ Fy \ \& \ \neg Exy)$' can be interpreted in T^1 using Quine's method as "There are in Geach's volume two non-identical type-words x and y and both involve involve two initial segments each of which ends with the letter 'e' " or in T as "There are in Geach's volume two non-equiform token-words x and y, and both words contain two letter 'e' ", 'the changes of import for the bits of the formula cancel out, and the truth condition of whole remains unchanged.'

(4) Given that the truth-conditions of any sentence belonging to T^1, T^2, T^3... are preserved in T, the ontological commitments of this sentence are also carried over to T because, as Geach (1973: 299) puts it, 'it is flatly inconsistent to say that as a member of a larger theory a sentence retains its truth conditions but not its ontological commitments'.

(5) Consequently, the application of Quine's method to the different sub-theories T^1, T^2, T^3... that are carved from T leads to a proliferation of the entities to which the sentences of T are committed to be true —a proliferation that is at odds with Quine's methodological conservatism in ontology.

If this argument is sound, it is clear that Quine's method for reducing identity is not a very attractive option to respond to the first horn of the dilemma set up by Geach. But matters are even worse for advocates of Quine's method according to Geach. Indeed, he also contends that the method is not merely at odds with Quine's methodological conservatism, but that it is deeply problematic because it leads to outright logical absurdities. To show this, Geach first introduces the predicate 'x is the same surman as y', which he defines as 'x is a man with the same surname as y, who is also a man'. Having done that, he considers a certain theory T with a universe of discourse that contains only men and a theory T' (which is a sub-theory of T) that contains the abovementioned predicate and that is unable to distinguish different men having the same surname. With these elements in play, he then introduces the following argument:

Suppose, for example, that we have in a sublanguage no predi-

> cables to distinguish two men with the same surname. Then if a theory T in the fragment of the main language is committed to the existence of men, the fragment of T in that language will be ontologically committed, *if construed in this way* [i.e., using Quine's method], not only to surmen, but of creatures for whom the predicate 'is the same surman as', as I defined it, supplies a creterior for absolute identity. Let us say, for short, on the suggestion we are discussing, this fragment of T is committed to the existence of *absolute surmen*. If so, then T itself is likewise committed, since a theory picks up the ontological commitments of each theory it contains. But the existence of absolute surmen, I shall argue, is an absurd supposition. (1973: 299-300)

As we can appreciate, the main idea underlying Geach's argument in this passage is that using Quine's method to interpret a sentence such as 'Theodore Roosevelt is the same surman as Franklin Delano Roosevelt' as expressing identity in the subtheory T' compels us to accept that, in addition to be committed to the existence of men, the former sentence is *also* committed to surmen, which are bizarre entities that seem to have contradictory properties. Indeed, on the one side, a surman is pretty much like a man insofar as he has a brain, a heart and many other features that characterize individual men (Geach 1967: 10) but, on the other side, a surman is unlike any individual man insofar as he has features that no man has. For instance, since Theodore Roosevelt and Franklin Delano Roosevelt are the same sur-

man, a surman can be in different locations at the same time.[20] In light of this, Geach's argument may be reconstructed in the following way:

> (1) Suppose that there are two theories, T and T', such that the universe of discourse of both includes only men and that, while T can distinguish between different men with the same surname, T' (which is a part of T) cannot.
>
> (2) Suppose that both T and T' contain the predicate 'x is the same surman as y' (which stands for 'x is a man with the same surname as y, who is also a man') and that, using Quine's method, we construe that predicate as expressing identity in T'.
>
> (3) The use of Quine's method on T' to ensure that the predicate 'x is the same surman as y' expresses identity forces us to modify the meaning of the remaining predicates of T' as well as the ranges of its quantifiers.
>
> (4) Granting that the quantified variables of a theory are the bearers of its ontological commitments, the use of Quine's method on T' forces us to acknowledge that T' is committed to the existence of surmen.
>
> (5) Since the ontological commitments of any sentence belonging a subtheory T' carry over to any theory T that contains it in virtue of set-theoretical inclusion, the sentences of T' that are

[20] In light of this, it is clear that surmen are objects that simultaneously belong to different ontological categories, which explains why Geach (1967: 10) maintains that the use of Quine's method makes the universe of discourse to adopt a 'baroque Meinongian structure'.

committed to surmen are also committed to the existence of surmen in the richer theory T.

As we can appreciate, this argument is different from the previous one insofar as Geach does not invoke here the sameness of truth-conditions of the sentences involving the I-predicate 'x is the same surman as y' in T' and T' to justify that the ontological commitements of the sentences involving the predicate in T' are carried over to the larger theory T of which T' is a fragment. Here, the key move consists in arguing that the ontological commitments of the sentences involving the I-predicate in T' and reinterpreted in order to express identity are preserved in T, not in virtue of the sameness of their truth-conditions in both theories, but rather in virtue of the properties of set-theoretical inclusion.

Thus, there seem to be in Geach's works at least three main arguments that are given to support RIT_1. The first argument, which we may call the Dilemma Argument, sets a dilemma that advocates of classical identity face and aims to show that neither alternative is satisfactory. The two other arguments, which are intended as refutations of the method suggested by Quine to ensure that at least within the context of a theory with a finite vocabulary a certain predicate expresses classical identity, intend to establish different things. One argument, which we may call the Truth-Conditions Argument, aims to show that using Quine's method leads to an ontological proliferation of entities, which violates Quine's methodological conservatism in ontology. The other argument, which we may call the Meinongian Argument, aims to establish that the kind of entities that the use of Quine's method commits us to are entities with contradictory properties that belong in a Meinongian

slum.

2.3.2 Geach's arguments for RIT_2

In the previous subsection, I have considered Geach's main arguments in favor of RIT_1, which is the core thesis that underlies RIT. In this section, I turn to Geach's arguments in favor of the two versions of RIT_2. As I mentioned earlier in §2.2.2, Geach stresses the connection between the notions of identity and cardinality and he appeals to Frege's view on attributions of number to justify the sortal relativity of identity in several passages. From these passages, one may extract the following argument in favor of RIT_2:

> (1) There is an intimate connection between the notion of identity and that of cardinality, which is illustrated by the fact that asserting certain identity statements is tantamount to assert certain corresponding cardinality statements.
>
> (2) In order to be meaningful, cardinality statements need to involve, as Frege showed, a certain sortal term that, in virtue of supplying both a criterion of identity and a criterion for counting, specifies what items are denoted by the statements in question.
>
> (3) Given the intimate connection between the notions of identity and cardinality, statements of identity also need to involve a certain sortal term that specifies what kind of identity holds between the items that the statements of identity refer to.

Though Geach appeals to Frege's view on cardinality attributions to build this argument, he also voices serious criticisms of Frege for failing to endorse

the thesis of sortal relativity of identity. To be more specific, since Frege endorses the view that meaningful statements involving attributions of cardinality require a sortal term but maintains also that identity is absolute, Geach censures him for not realizing that these views are at odds with each other and that, given his views on cardinality, he should have endorsed the thesis of sortal relativity of identity.[21]

In addition to the aforementioned argument (which we may refer to as the Cardinality Argument insofar as it exploits the intimate connection between identity and cardinality), Geach also presents another line of reasoning to establish the thesis of sortal relativity of identity. This line of reasoning consists in showing that RIT_2 is a very attractive thesis insofar as it provides a satisfactory solution to several puzzles. To prove this, Geach presents a puzzle (which he attributes to William of Sherwood) that is respresentative of the type of problems that RIT_2 enables us to deal with. According to the puzzle, if we consider a certain cat named 'Tibbles' that is sitting on a mat and if we assume that it has at least 1000 hairs (which are numbered h_1, h_2,... h_{1000}), it would seem that, contrary to our initial stipulation, we are compelled to accept that there are 1001 cats sitting on the mat in virtue of the following reasoning:

> Now let c be the largest continuous mass of feline tissue on the mat. Then for any of our 1,000 hairs, say h_n, there is a proper

[21] I have previously distinguished two versions of the sortal relativity thesis —namely, RIT_{2m} and RIT_{2e}, so one might raise the following question: which version of RIT_2 does the previous argument support? I am inclined to maintain that, if it does indeed support the thesis of sortal relativity of identity, it provides better support to RIT_{2e} because the argument appeals to Frege's view and Frege maintains that the notion of criterion of identity has an epistemic function rather than a metaphysical one.

> part c_n of c which contains precisely all of c except that hair h_n: and every such part c_n differs in a describable way both from any other such part say c_m and from c as a whole. Moreover, fuzzy as the concept *cat* may be, it is clear that not only is c a cat, but also any part c_n is a cat: c_n would clearly be a cat were the hair h_n, to be plucked out, and we cannot reasonably suppose that plucking out a hair generates a cat, so c_n must already have been a cat. (1980: 215)

Indeed, given that the same point that Geach makes vis-à-vis c_n can also be made vis-à-vis any proper part of Tibbles that differs from it by lacking one hair at most, it would seem that, if we accept that c_n is a cat, we have to accept that any proper part of Tibbles that differs from it just by one hair is also a cat. And this apparently forces us to acknowledge that there are 1001 cats sitting on the mat —i.e., c and all its proper parts that differ from it by lacking one hair. Of course, this conclusion is intolerable for Geach, and he accordingly seeks to provide a satisfactory way to solve the puzzle.

The solution that he advocates consists in invoking the thesis of sortal relativity of identity. Indeed, if we assume (as he does) that the predicate 'x is the same cat as y' expresses a relative identity relation which holds between c and any of its proper parts c_1, c_2...c_{1000}, we can solve the puzzle by arguing that, even though the proper parts c_1, c_2...c_{1000} are all different *masses of feline tissue*, they are nevertheless one and the same *cat*. Thus, appealing to the sortal relativity of identity allows us to preserve the commonsensical claim that there is only one cat one the mat while also accounting for the idea that, in a sense, c and its proper parts c_1, c_2...c_{1000} are all different entities.

In light of this, we may offer the following reconstruction of Geach's second argument (which we might call the Puzzle-Solving Argument):

> (1) There are puzzles that have the following form: we consider a certain entity e belonging to a certain kind K and, since we acknowledge that e has many proper parts e_1, e_2...e_n such that each one of them would be e if some minimal changes took place, we are forced to accept that there are many other entities belonging to the kind K that coincide with e.
>
> (2) This coincidence of multiple entities of the same kind K in a single spatio-temporal region is intolerable insofar as it violates a commonsensical principle enunciated by Locke according to which no two things of the same kind K can be in the same place at the same time.[22]
>
> (3) A solution to these puzzles that enables us both to preserve Locke's principle and to acknowledge that e and its proper parts e_1, e_2...e_n are, in a sense, different entities consists in appealing to the thesis of sortal relativity of identity.
>
> (4) There are other solutions to this these puzzles, but they are far less attractive because they entail either a violation of Locke's principle or rejecting the claim that the proper parts e_1, e_2...e_n of e also belong to the same kind K as e.
>
> (5) Consequently, there is a good reason to accept the thesis of

[22]Locke (1995: 242) formulates the principle in the following way: '[the principle] is existence itself, which determines a being of any sort to a particular time and place incommunicable to two beings of the same kind.'

sortal relativity of identity insofar as it provides the best available solution to certain puzzles.[23]

As we can appreciate, Geach's case for RIT_2 primarily relies on two arguments, the Cardinality Argument and the Puzzle-Solving Argument, which seem to be aimed respectively at supporting RIT_{2e} and RIT_{2m}. Having made clear what are the main arguments deployed to support the two version of RIT_2, I want to consider now the connection between these two arguments and other minor theses that Geach endorses as part of RIT. A thorough look at the second premise of the Cardinality Argument allows us to see that the notion of sortal term, which provides us both with a criterion for counting and a criterion of identity, plays a crucial role in Geach's reasoning. Indeed, without sortal terms, we are unable to determine whether a statement that involves a cardinality assertion is true or not because we have no way of distinguishing items that have been counted already from items that have not been counted yet. Thus, if we grant that counting items of a certain kind K involves being able to put them into a one-to-one correspondence with the natural numbers,[24] it is clear that our ability to put items into a one-to-one correspondence with the natural numbers presupposes the ability to distinguish them. Now, one may distinguish and count items on the basis of classical identity. But, if Geach is right and there is no way to express

[23]Just as we raised in footnote 21 a question concerning which version of RIT_2 the Cardinality Argument aims to bolster, we may also ask which version of RIT_2 the Puzzle-Soving Argument intends to provide support to. Considering that the puzzle that Geach presents is metaphysical in nature (i.e., what worries Geach is avoiding the absurd thesis that there are 1001 cats on the mat), I believe that the Puzzle-Solving Argument aims to provide support to RIT_{2m} rather than to RIT_{2e}.

[24]For a recent defense of this claim, see Heck (2000). See also, for the original statement of this claim, Frege (1894: 319).

classical identity, we may still distinguish and count items using other equivalence relations weaker that classical identity. This point is captured by the following thesis that Geach also endorses as part of RIT:

> (RIT_4) In the absence of classical identity, one can distinguish and count items using weaker equivalence relations without any inconsistencies or contradictions.

In order to appreciate this, consider again Geach's example involving the words contained in a volume in one of his shelves. One may distinguish and count token-words using numerical identity, but one may also distinguish and count the token-words using a weaker equivalence relation, such as the relation *is located on the same page as*. If we use this relation to distinguish and count token-words, all the words located on a certain page are counted as one. Though this way of distinguishing and counting might seem unusual, there are no inconsistencies or contradictions associated with it. In light of this, we can appreciate that RIT_2 (in either the metaphysical or the epistemological version) provides an important motivation to endorse RIT_4.

In addition to provide motivation for RIT_4, any of the two versions of the thesis of sortal relativity of identity entails another thesis that Geach also endorses as part of RIT, which is the irreducibility of restricted quantification. In fact, the irreducibility of restricted quantification is a direct consequence of the claim, which is part of both RIT_{2m} and RIT_{2e}, that relative identity predicates are primitive. Indeed, if one cannot analyze the relative identity predicate 'x is the same river as y' as 'x is an river, y is an river and x is the same as y', it follows, according to Geach, that one must take instances of

restricted quantification (e.g., quantification over *rivers*) as being irreducible to unrestricted quantification (i.e., quantification over *things* that are rivers).

To show this, Geach (1962: 150-151) asks us to consider the true sentence 'Heraclitus bathed in some river today, and bathed in the same river yesterday' and argues that, even if we accept that every river is a mass of water (which is a plausible assumption) and an analysis of the previous sentence as 'For some x that is a river, Heraclitus bathed in x today, and bathed on the same x yesterday', one cannot derive from the assumption and the proposed analysis the sentence 'For some x that is a mass of water, Heraclitus bathed in x today, and bathed in the same x yesterday'. Indeed, since the waters of any river are in constant flow, it may turn out to be false that Heraclitus bathed in the same mass of water today as he did yesterday (unless he traveled downstream) even if it is true that he bathed in the same river in two successive days. It is in virtue of this type of problem that Geach maintains that a sentence that involves restricted quantification such as 'Some rivers are long' cannot be analyzed in terms of unrestricted quantification (i.e., it cannot be analyzed as 'There are some things that are rivers and long'). This thesis, which is a minor element of RIT, may be formulated as follows:

> (RIT_5) Restricted quantification over Fs (where F is a sortal term) is more basic than unrestricted quantification over *things* that are F and, consequently, the former cannot be analyzed in terms of the latter.

It is important to notice that Geach does not reject unrestricted quantification as inconsistent or contradictory. He clearly allows unrestricted

quantification to the extent that he admits that one may obtain the one-place predicate 'x is the same F as something' by derelativization from the relative identity predicate 'x is the same F as y'. His claim is just that restricted quantification is more basic than unrestricted quantification and that, accordingly, the former cannot be understood in terms of the latter.

The distinction between unrestricted and restricted quantification as Geach presents it is also important as it might be used to illustrate a further minor thesis that Geach endorses, which concerns the use of names. According to Geach, a name such as 'Jemima' may be used to designate a *cat*, but it may also be used to designate a *thing* that is a cat. In the first case, considering that the name 'Jemima' is use to refer to an entity in an restricted domain of quantification that involves only cats, it is, as Geach puts it, a name *for* a cat whereas in the second case, where the name is used to refer to an entity in an unrestricted domain of quantification, it is a name *of* a cat. This thesis may be formulated in the following way:

> (RIT_6) A name can be used to refer to an F in a restricted domain of quantification (in which case it is a name *for* an F) or to refer to a thing that is F in an unrestricted domain of quantification (in which case it is a name *of* an F).

As Noonan observes, Geach uses the distinction between a name *for* an F and a name *of* an F to present an account of the truth-conditions of sentences involving instances of unrestricted quantification and illustrate their relation to the truth-conditions of sentences involving instances of unrestricted quantification. Here is how he proceeds: after remarking that there are cases in

which one cannot provide truth-conditions for sentences of the form 'F(some A)' or 'F(any A)' by listing the As (in particular, cases in which the As are too many or in which the the group of As is open toward the future), Geach (1980: 206) puts forward the following suggestions for both cases:

> 'F(some A)' is true iff '$F(a)$' is true for some interpretation of 'a' as a name *for* and *of* an A.
>
> 'F(any A)' is true iff '$F(a)$' is true for any interpretation of 'a' as a name *for* and *of* an A.

Having presented the aformentioned equivalences as providing the truth-conditions for sentences involving instances of restricted quantification over Fs, Geach then remarks that if we remove the restriction to proper names *for* and *of* Fs, we obtain truth-conditions for sentences of the form 'For some x, $F(x)$' and 'For any x, $F(x)$', which involve unrestricted quantification. Thus, the distinction between a name for an F and a name of an A enables Geach to show that truth-conditions for sentences involving unrestricted quantification can be obtained from truth-conditions for sentences involving restricted quantification by eliminating the restrictions.

2.3.3 Geach's argument for $\mathbf{RIT}_3$

Finally, we may turn to consider the argument that Geach provides to make a case for the derelativization thesis. The main argument that Geach gives to defend RIT_3 consists in showing that every term that we characterize as a sortal (i.e., a term which supplies both a criterion to count and a criterion of identity for entities of a certain kind K) must be obtained from a certain

relative identity predicate by derelativization. Indeed, for him, this is the only non-circular way to explain the relation between the expressions 'x is the same A as y' and the sortal term 'x is an A' as the following passage illustrates:

> We can offer 'is the same A as *something*' to analyze away the predicative use, 'is an A', of a substantival general term. For our using the expression 'is the same A as' requires 'A' to be construed as such a term, and in that case a thing is an A iff it is the same A as something or other; and there is no risk of a vicious circle, just because 'is the same A as' does not admit of the analysis 'is an A and is the same as', which would leave 'is an A' again on our hands. (1980: 213)

According to Geach, we cannot explain the relation between the expressions 'x is the same A as y' and 'x is an A' by analyzing the former in terms of the latter because doing so lands us in a vicious circularity. This cicularity is clearly illustrated for Geach by the fact that, after offering the expression 'x is an A, y is an A and x is the same as y' as a tentative analysis of the relative identity predicate 'x is the same A as y', one may still consistently ask the question 'x is the same *what* as y?' and that the answer is nothing but the original relative identity predicate that we aimed to reduce. Thus, the argument in favor of RIT_3 may be formulated in the following way:

> (1) There are two options to account for the relation between a relative identity predicate of the form 'x is the same A as y' and its corresponding sortal predicate 'x is an A': one may obtain

the relative identity predicate by restricting classical identity to a certain domain given by the extension of the sortal predicate 'x is an F' (which is taken as primitive) or one may obtain the sortal predicate 'x is an A' by derelativization of the relative identity predicate 'x is the same A as y (which is taken as primitive).

(2) Out of these two options, the first one is problematic because it fails to provide a non-circular account of the relation between the relative identity predicate and its corresponding sortal.

(3) Accordingly, the best option we have to account for the relation between a relative identity predicate and its corresponding sortal is by means of the method of derelativization.

Geach provides futher support for this argument be presenting a series of examples that illustrate how we may successfully obtain sortal terms by applying the method of derelativization on certain relative identity predicates. Now, after presenting the main argument that Geach provides to support the main theses that make up, we may turn now to a critical discussion of them.

2.4 Some traditional objections against RIT

In this section, I present some traditional objections to Geach's arguments in support the three main theses that constitute the core of RIT. Following the order of exposition that I adopted in the previous section, I will rehearse first the objections raised against RIT_1, and I will subsequently review the criticism addressed to both RIT_2 and RIT_3.

2.4.1 Some traditional replies against $\mathbf{RIT}_1$

As we previously saw, there are three main arguments that Geach deploys in order to make a case for RIT_1 —namely, the Dilemma Argument, the Truth-Conditions Argument and the Meinongian Argument. Out of these three arguments, the Truth-Conditions Argument and the Meinongian Argument have elicited the strongest interest in the philosophical literature. The main reason for this interest lies partially on the structure of Geach's reasoning. Indeed, given that the Dilemma Argument consists in setting up a dilemma for advocates of classical identity and that Geach maintains that adopting of one of the horns of the dilemma is precluded by the Truth-Conditions argument and the Meinongian Argument (which are directed against the use of Quine's method) while adopting the other is precluded in virtue of certain semantic paradoxes, one can defuse the Dilemma Argument in principle by showing that both the Truth-Conditions Argument and the Meinongian Argument are flawed. In light of this, I will present in the following two subsubsections the objections that have been addressed to these two arguments (and I will add some objections of my own).

2.4.1.1 The Truth-Conditions Argument

As we have previously shown, the Truth-Conditions Argument crucially relies on the thesis that, if we apply Quine's method to reinterpret the sentences of a certain subtheory T' that involve an I-predicate in order to ensure that this predicate expresses identity, the sentences have, in a larger theory T of which T' is a part, not only the same truth-conditions but also the same ontological commitments that they have in T'. Noonan has vigorously

criticized this premise of the argument, arguing that it does not correspond to the notion of ontological commitment defended by Quine:

> For Quine, the ontological commitments of a theory are those entities which must lie within the domain of quantification of the theory if the theory is to be true; or, alternatively expressed, those entities the predicates of the theory have to be true of if the theory is to be true. A theory is not, if I may so express it, ontologically committed to what is required to be in *the* universe if it is to be true, but merely to what it is required to be in *its* universe if it is to be true. (1997: 644)

Thus, for Noonan, the Truth-Conditions Argument founders because Geach fails to appreciate that, when he considers the truth-conditions of a sentence s of T' involving an I-predicate that has been reinterpreted (along with the rest of the lexicon and the quantifiers of T') using Quine's method so that it expresses identity, these truth-conditions for s fail to be preserved in a larger theory T of which T' is a part because the universe of discourse of T is different from that of T'. Similar objections have also been raised by other critics. In particular, Nelson has also convincingly argued that there is no preservation of the ontological commitments of a sentence s of T' in a larger theory T of which T' is a part because, when we move from T' to T, the meanings of the predicates of and other lexical elements of T' are not preserved in T:

> If we start, in T, with a domain of token-words, and if we move to T' by omitting the predicables that allow us to discern different

> token of the same type word, then, to be sure, a revision in out ontology is called for. The obvious revision is to substitute an ontology of type words for our old ontology of token-words. So doing will give us a new ontology, an ontology contianing different things than did our old ontology. But it is hard to see what sense there is in the claim that an ontology of type words is larger than an ontology of token words (1970: 252)

Immediately after this passage, Nelson considers the claim, which is presumably what Geach has in mind, that the use of Quine's proposal to ensure that an I-predicate belonging to a theory T' expresses identity forces us to admit not only that true sentences of T' involving the I-predicate are committed to certain entities but that their ontological commitments are carried over to a larger theory T of which T' is a part. Nelson maintains that Quine would reject such a conclusion on the grounds of his ontological conservatism, but he does not show explicitly why Quine is not forced to accept the ontological inflation. To bolster the case that both Noonan and Nelson mount against the Truth-Conditions Argument, I want to offer now an objection of my own that shows clearly why the use of Quine's method does not force us to accept the kind of out-of-control ontological proliferation that worries Geach.

As we have previously seen, the use of Quine's method to ensure that an I-predicate expresses identity in a certain theory T' involves a reinterpretation of the lexicon of T' and a modification of the range of the quantifiers of T'. Now, the purpose of these changes is to offer a *model* in which true sentences of T' that contain the I-predicate in question are reinterpreted as

being true and as expressing identity. If we consider Quine's method as an attempt to provide a model with certain characteristics for true sentences of T' that involve an I-predicate, Geach's objection to Quine can then be recast in the following way: since Quine's method forces us to acknowledge certain models that make true certain sentences of a theory T' (specifically, sentences that include an I-predicate construed as expressing identity), these models also make true the sentences in question when they are considered in the framework of a larger theory T of which T' is a part. Thus, in addition to being committed to certain entities that constitute a model that makes true its sentences, a theory T is also committed according to Geach to the entities that constitute all the models for all the sub-theories that are part of T if Quine's method is applied to reconstrue the sentences involving I-predicates contained in these sub-theories.

But this cannot be right. Even if a sentence admits different models that make it true, this does not entail that it is ontologically committed to all the entities that make up these models. Intuitively, certain models just seem to be better than others. But what principle could we rely on to justify this intuitive claim? In the context of a discussion on Quine's views on ontological reduction, Grandy suggests what seems to me an extremely plausible way to distinguish wanton models for a certain theory from those that are really called for to make its sentences true:

> The general idea is that, when we are giving an interpretation of a [theory], it is otiose to have objects in the model which are unnecessary. Thus, for example, a model with undistinguishable objects or one in which a certain subset of the domain could be

> dropped should not be seriously contemplated as interpretation of the [theory]. (1979: 71)

If we rely on the suggestion that Grandy presents in this passage (which is an application of Ockham's razor), there is an effective way to resist Geach's claim that the use of Quine's method entails a distasteful proliferation of entities. Indeed, even though one might concede that a certain model $\mathcal{M}'$ that involves the existence of the entities that Geach calls surmen makes the sentence 'Theodore Roosevelt is the same surman as Franklin Delano Roosevelt' true not only in a theory T' in which the I-predicate 'x is the same surman as y' is construed as expressing identity but also in a richer theory T (which contains T') that allows us to distinguish different men with same surname, this does not entail that T is committed to a surman that is both Theodore Roosevelt and Franklin Delano Roosevelt. Considering that the expressive ressources of T allow us to distinguish different men with the same surname, all that we require to make the sentence true in T is a model $\mathcal{M}'$ that contains men.

2.4.1.2 The Meinongian Argument

Though it is connected to the Truth-Conditions Argument that Geach deploys against RIT_1, the Meinongian Argument has a different focus as I noticed at the end of §2.3.1. Indeed, while the Truth-Conditions Argument intends to show that the use of Quine's method violates his methodological conservatism in ontology, the Meinongian Argument aims to establish that Quine's method commits him to entities that have contradictory properties —e.g., surmen which have brains and hearts like ordinary men but which

can also be present at different places. In response to this argument, several philosophers have maintained that the use of Quine's method does not commit us to any Meinongian entities, *pace* Geach. In particular, Perry has argued that, even though Geach maintains that in order to construe the I-predicate 'x is the same surman as y' as expressing identity in certain theory T' to who Quine's method is applied one is forced to reinterpret the lexicon of T' in such a way that we end up committed to surmen —which Geach (1967: 10) describes as 'androids who differ from men just in in this respect, that two different ones cannot share the same surname', the reinterpretation of the lexicon of T' does not force us to adopt a Meinongian ontology:

> As far as I can see, nothing more objectionable than families would emerge from this reinterpretation. I cannot see why Geach thinks it should require androids. The entity that has all the persons with a certain last name as occurrences (parts or members) is clearly something like a family and not anything like an android. (1970: 196)

A similar reply to the Meinongian Argument is put forward by Nelson who maintains that, though the use of Quine's method to ensure that an I-predicate in T' expresses identity forces us to modify the range of quantifiers of T', this modification does not force us to adopt queer and controversial entities. For Nelson (1970: 253), 'the plausible Quinean move would be to make the ontology of $[T']$ an ontology of classes, classes of men with the same surname.' In light of this, maintaining that surmen are like men reveals a serious confusion in Geach's thought because he treats entities that clearly belong to one ontological category —namely, classes— as if they were part

of another ontological category —namely, men. It is this confusion which gives the impression that the use of Quine's method commits us to bizarre entities.

In response to Perry and Nelson, Geach (1973: 295) maintains that, since he stipulates that the predicate 'x is the same surman as y' means 'x and y are men and have the same surname', he intends the predicate to exclusively hold between surmen and not between families or classes of men, and he also maintains that, insofar as 'nobody has shown any incoherence in the definition [of the predicate 'x is the same surman as y'] as calls for conjectual emendation of my text to restore sense', he is then perfectly free to define the predicate as being true of only of surmen (which are men), and not of classes of men or families.

However, even if Geach is *prima facie* correct when he maintains that he can define the predicate 'x is the same surman as y' as he wants (e.g., as applying to only to surmen), he owes us an explanation of the characteristics that surmen have beyond the mere stipulation he makes. This is due to to the fact that, as Robinson (1950: 78) has remarked when reflecting on the habit of making stipulation in philosophy, 'the habit of stipulating one's own meaning for words tends to bring with it the habit of evading the analysis of obscure conceptions and the clarification of actual meanings'. To be more specific, since every surman is a man according to Geach, one should be entitled to ask him how tall or how old is a specific surman (say, the surman that is both George Walker Bush and John Ellis Bush)[25] or where he is located, since having a certain height, a certain age and a certain location

[25]In fact, each male member of the Bush family is the same surman as George Walker and John Ellis, but I ignore this issue here for the sake of simplicity.

are usual characteristics that men have.

However, this poses a serious difficulty for Geach. Assuming that we denote the surman that both George Walker Bush and John Ellis Bush are as 'Geojeb' and that George is 64 years old while John is 57, we can raise a question about the age of Geojeb. But Geach then faces a dilemma: on one side, if he provides a specific numerical answer to the question, he is liable to the charge that this answer is arbitrary (indeed, there seem to be no criteria to determine Geojeb's age) and, on the other side, if he answers that Geojeb either has no age or that he is both 64 and 57 years old, his contention that every surman is a man is deprived of its substance. Thus, since Geach's response to the objection raised Perry and Nelson fails, it is clear that the Meinongian Argument founders.

Now, both Nelson and Perry agree on the fact that the use of Quine's method to ensure that an I-predicate expresses identity does not commit us to bizarre Meinongian entities. However, though both claim that the use of Quine's method only commit us to uncontroversial entities such as classes of men or families, they do not offer a reason that justifies the claim that the use of Quine's method does not commit us to Meinongian entities other than claiming that these entities can be safely dispensed with. Before turning our attention to the arguments in favor of RIT_2, I want to present here a more substantial reason drawn from the analysis of Quine's notion of ontological commitment to argue that, *pace* Geach, the use of Quine's method does not commit us to a universe of discourse with a baroque Meinongian structure.

In order to do this, let me first rehearse an important point that Cartwright made concerning Quine's notion of ontological commitment in one of his ear-

liest papers. After analyzing several passages drawn from Quine's works in which the notion of ontological commitment is discussed, Cartwright (1954: 319) remarks that one common element in all these formulations is that 'each formulation specifies the ontology of a theory as consisting of just those objects over which the bound variables *must* range, if the affirmed statements of the theory are to be true.' In light of this, Cartwright argues that Quine's notion of ontological commitment is underpinned by a series of intensional notions.

Taking Cartwright's remarks as inspiration, Church argues subsequently that, not only does the notion of ontological commitment presuppose intensional notions, but that ontological commitment is 'an intensional notion, in the sense that ontological commitment must be to a class concept rather than a class' (1958: 1013). The main rationale to endorse this view is that, if we argue that ontological commitment is to individual items rather than to concept-classes, a theory that contains the sentence 'Unicorns do not exist' would have to be committed to unicorns to make that sentence true —which is absurd. Instead, if ontological commitment is to concept-classes, the theory only has to be committed to the existence of the concept *being a unicorn* and to its being subsumed under the concept *having no instances* to make the sentence true.

Since the view endorsed by Church presents a clear advantage when dealing with sentences that refer to alleged non-existent objects, some authors in addition to Church have also endorsed it. In particular, when Jubien considers a sentence T that requires the existence of an object a to be true, he argues that, in a case where the singular term 'a' fails to designate any ex-

tensional object, a dilemma arises which clearly illustrates that the notion of ontological commitment is intensional in the sense that Church characterizes:

> Either we permit the second position of the context ['a'] to range over intensional objects, such as concepts or whatever, or else we find a way of interpreting 'T assumes a' which makes a suitable accommodation for the cases in which 'a' fails to designate. (1972: 384)

Given that Jubien subsequently shows that there is no plausible way to interpret the sentence 'T assumes a' that accommodates the cases in which 'a' fails to refer, he ultimately concludes that ontological entailment really commits us to concepts. And, if this is indeed the case, we have a good reason to argue that the use of Quine's method does not commit us to Meinongian entities. Indeed, assuming that Church and Jubien are right, what Geach's Meinongian Argument shows at best is that, if a theory T contains the sentence 'George is the same surman as Jeb' and if Quine's method is used on T to make the predicate 'x is the same surman as y' express identity, T is not committed to any actual surmen to make the sentence true but merely to the existence of the concept *being the same surman as* and to the fact certain entities, which are men, fall collectively under it under it. Consequently, if the notion of ontological commitment that Quine endorses is intensional as both Church and Jubien argue, Geach's Meinongian Argument clearly fails to undermine the use of Quine's method.

2.4.2 Some traditional replies to RIT_2

Having considered in the former subsection certain replies to some of the arguments that Geach deploys to argue in favor of RIT_1, we can turn now to a critical assessment of the arguments he provides to argue for RIT_2. Considering that there are two versions of RIT_2 and that I have argued that each version is supported by a specific argument, I will first consider the Cardinality Argument, which aims to support RIT_{2e} and, having done that, I will go over the Puzzle-Solving Argument, which aims to bolster RIT_{2m}.

2.4.2.1 The Cardinality Argument

As a thourough analysis of its structure shows, the Cardinality Argument crucially relies on the premise which expresses the existence of an intimate conncetion between the notion of identity and that of cardinality. Now, I do not want to dispute the fact that accepting certain identity (or non-identity statements) entails accepting certain cardinality statements: it is clear that, if you acknowledge that x is not identical to y, you are committed to the fact that x and y are (in some sense) two. However, what I do want to argue against is the use that Geach makes of this connection. In particular, I want to show, following the footsteps of both Perry (1970) and Blanchette (1999), that Geach's appeal to Frege's view on cardinality to provide support for RIT_{2e} betrays some serious misunderstandings of Frege's position to which Geach is prey.

In order to appreciate this clearly, let us first point out the similarities bewtween Frege's proposal concerning cardinality attributions and Geach's proposal regarding identity attributions. A common element to both propos-

als, which is highlighted by Perry (1970: 182-183), consists in the fact that sortal terms make a fundamental contribution to the assignment of truth-conditions to statements invovling either cardinality attributions or identity attributions. Indeed, if there is no sortal term F associated (either explicitly or implicitly) with the cardinality statement 'a and b are two', we are unable, according to Frege, to determine if the statement is true or false insofar as we have still to determine which are the items counted. Similarly, Geach maintains that, if there is no sortal term associated (either explicitly or implicitly) with the identity statement 'a is the same as b', we are unable to determine whether the aforementioned identity statement is true or false insfoar as we still have to determine which are the items identified. Thus, sortal terms play a key role in allowing us to formulate meaningful cardinality statements for Frege and meaningful identity statements for Geach.

A second feature common to Frege and Geach is, as Blanchette (1999: 206) remarks, that both authors maintain that cardinality attributions are not directly made vis-à-vis objects or vis-à-vis collections or aggregates. According to Frege, there is no specific number that corresponds to a given collection of items given that no collection of items presents us with strictly given units, as the fact that one and the same collection may correctly counted 'as either one poem, or as 24 books, or as some large number of verses' illustrates. Similarly, Geach maintains that is does not make sense to attribute cardinality to a certain collection of items; cardinality attributions are only meaningful relative to a certain sortal term that specifies which items are to be counted. Thus, the second similarity between Frege and Geach consists in the fact that both reject the thesis that the subject of cardinality attributions

are collections.

However, even if there are certain similarities between Frege's view on cardinality and Geach's view on identity, these similarities are not strong enough to warrant the use of Frege's view to provide support for the thesis of sortal relativity of identity, *pace* Geach. Indeed, even though Frege certainly maintains that sortal terms play a key role in cardinality statements, the role that he attributes to them in cardinality statements is different from the role that Geach attributes to them in identity statements. For Geach, the purpose of a sortal term involved in an identity statement is to determine which specific identity relation is involved there. For Frege, the role of a sortal term in a cardinality statement is not to determine which specific type of numbers are used to count but rather, as Perry (1970: 185) remarks, to identify the kind of items that are counted. Thus, even if Frege and Geach accept that sortal terms play a key role in both cardinality and identity statements, this agreement concerns a minor issue and masks a profound divergence in their views —a divergence that undermines Geach's attempt to use Frege's view to support the thesis of sortal relativity of identity.

Now, when we consider the fact that both Frege and Geach hold that cardinality attributions are not made vis-à-vis collections or objects, we can appreciate that there are key differences in the reasons that use to defend this thesis. According to Frege, it does not make sense to attribute cardinality to a collection or to an object because there is no number that corresponds specifically to a collection or an object. Any collection or object can be numbered properly in many distinct ways because cardinality attributions are made, not vis-à-vis the collection or the object in question, but vis-

à-vis different *concepts* under which the relevant collection or object falls. For Geach, on the contrary, cardinality attributions cannot be done vis-à-vis collections or objects *simpliciter*; however, they can be made vis-à-vis collections or objects under a sortal F.[26] Thus, since the motivations for endorsing the thesis that cardinality attributions cannot be made vis-à-vis collections or objects are radically different in both authors, Geach cannot appeal to the fact that Frege endorses this thesis to provide support for the sortal relativity of identity.

Considering that the similarities between Frege and Geach about the role of sortal terms and the subject of cardinality attributions are rather superficial, Geach has no solid reason to appeal to Frege to substantiate his thesis. But his appeal to Frege's views on cardinality is not only unsubstantiated; it is problematic insofar as it relies on a total misunderstanding of Frege. Let me explain. Though Geach distinguishes two different senses of the notion of identity criterion (which respectively correspond, as we have seen, to two different construals of the thesis of sortal relativity of identity), these two senses of the notion of identity criterion are underpinned by a single notion of sortal concept that Geach characterizes as follows:

> Having a concept never means being able to recognize some feature we have found in direct experience; the mind makes concepts, and this concept formation and the subsequent use of the concepts formed never is mere recognition or finding (...) (1957, 40)

[26]The reason for this is that, strictly speaking, there are no objects or collections *simpliciter* for Geach, as he makes clear when he (1973: 290) asserts that 'talk of bare particulars (...) is manifest nonsense'.

As this passage clearly shows, concepts have two primary characteristics for Geach: (1) though they play a role in allowing us to recognize or identify entities as belonging to a certain kind K, they are not merely circumscribed to this role and (2) they are not objective entities but rather mind-dependent. Now, if concepts do not merely have a role as 'markers' that enable us to identify things, they presumably also function as 'cookie-cutters' that enable us to carve entities out of some undetermined reality —which is the way that Dummett construes Geach's proposal. As I previously mentioned, this view of concepts (and the concomitant view of reality as an amorphous and undetermined blob that it entails) is extremely problematic because it is at odds with other theses endorsed by Geach elsewhere. In particular, Geach (1961) maintains, following Aquinas, that there is one determinate reality, which is God. Thus, the consequences of the view of concepts that Geach presents are clearly opposed to the realist position that he adopts —a fact which casts serious doubts on the consistency of Geach's general views.

But matters are even worse for Geach. Indeed, the view of concepts that Geach endorses is not only at odds with the realist position that he adopts, but it is also diametrically opposed to Frege's. Indeed, whereas Geach maintains that concepts are mind-made creations, Frege (1979: 37) clearly emphasizes that concepts are mind-independent, objective entities when he writes that 'as I do not create a tree by looking at it or cause a pencil to come into existence by taking a hold of it, neither do I generate a thought by thinking.' Since concepts are mind-independent entities according to Frege, Geach cannot appeal consistently to his view to substantiate the thesis of sortal relativity of identity. Indeed, according to Frege, for any cardinality

statement there is a fact of the matter concerning whether the statement is true or false, even if that fact is unbeknownst to us in virtue our cognitive limitations to grasp certain concepts. However, for Geach, if there had been no minds capable of creating and enterataining concepts, there would be no fact of the matter concerning the truth or falsity of any identity statement. In light of this, it is clear that Geach's attempt to derive support from Frege's view on cardinality for the thesis of sortal relativity of identity is deeply problematic insofar as it is ignores a fundamental difference between both authors: whereas Frege admits the existence of mind-independent cardinality facts, Geach rejects the existence of mind-independent identity facts.

2.4.2.2 The Puzzle-Solving Argument

Let us turn our attention to the Puzzle-Solving Argument. As a brief look at its structure shows, the key premise on which the argument depends is the claim that there are other solutions to the puzzle of 1001 cats, but that these solutions are far more onerous and problematic than the thesis of sortal relativity of identity. For Geach, we may solve the puzzle either by accepting the view that there can be more that one object belonging to a kind K occupying a certain spatio-temporal region or by rejecting the view that a proper part of a material object o that differs from it in only a small respect (say, one hair) is also a material object of the same kind as o, but any of these solutions involves difficulties. What I aim to do here is to show that, although these solutions seem to involve problems, Geach's argument for the implausibility of this thesis involves some very problematic assumptions and also that Geach himself provides in certain passages of his works suggestions

concerning possible solutions to the puzzle that do not involve endorsing the thesis of sortal relativity of identity.

As Geach presents things, the thesis of sortal relativity of identity is the best available solution to the puzzle insofar as the other solutions — i.e., accepting the existence of many other cats that overlap with Tibbles or denying that any of the proper parts of Tibbles that differ form it in just one hair are cats— are very problematic. Now, I do agree with Geach on the fact that the thesis that there are multiple overlapping cats that occupy almost the same region as Tibbles is implausible, if not absurd. But I do not believe that Geach mounts a solid case to show the implausibility of the second solution because his reasoning trades on a controversial assumption that has been questioned by Kripke.

In order to appreciate this, let us unpack Geach's argument in detail. Geach introduces initially the expression 'c_n' to refer in the actual world to Tibbles-minus-h_n (i.e., the proper part of Tibbles that includes all of it except h_n). Having done this, he considers what this expression designates in a world at which h_n has been plucked out and he argues that it refers to Tibbles. On the basis of this, he then claims that, since plucking hairs does not generate cats, what 'c_n' designates in the actual world (i.e., Tibbles-minus-h_n) must have already been a cat. As we can appreciate, the controversial assumption that Geach relies on in this reasoning is that, after using 'c_n' to fix a referent in the actual world, he can also use the property associated to that expression as providing a means to identify the original referent across counterfactual situations. But this assumption is extremely problematic, as Kripke (1980: 55-57) has shown in his discussion of the standard meter case.

Let me summarize the gist of Kripke's views. Kripke introduces a distinction between two roles that an expression may have: it can be used *to provide a synonym for another expression* or to *fix a referent for another expression*. These two roles must be clearly distinguished according to Kripke because a failure to do so yields serious problems, such as the difficulty explain how a statement such as 'The length of stick S at t_0 is one meter' can express a definition and nevertheless be contingent. If we embrace the distinction that Kripke suggests, the difficulty dissolves because:

> There is no difficulty between the counterfactual statement [that if circumstances had been differenct, the stick S at t_0 would not have been one meter] and the definition of 'one meter' as 'the length of S at t_0 (1980: 57)' because the 'definition', properly interpreted, does not say that the phrase 'one meter' is to be synonymous (even when talking about counterfactual situations), with the phrase 'the length of S at t_0 but rather that we have determined the reference of the phrase 'one meter' by stipulating that 'one meter' is to be a rigid designator of the length which is in fact the length of S at t_0 (1980: 56)

Since Geach uses 'c_n' to fix a certain reference in the actual world (namely, Tibbles-minus-h_n) and he subsequently considers the contingent property associated with the description that 'c_n' abbreviates (namely, the property *being the proper part of Tibbles that contains all of it except* h_n) as a means to identify the referent he picked in the actual world across different counterfactual situations, he implicitly assumes that the description enables him to pick the same object in counterfactual situations that he picked in the

actual world. But this is not the case, since the expression 'the proper part of Tibbles that contains all of it except h_n' is not a rigid designator: it designates Tibbles-minus-h_n in the actual world and Tibbles in the counterfactual world that Geach considers. In light of this, Geach is wrong to assume that he can use 'c_n' to pick a certain object in the actual world and then use the contingent property associated with the description that 'c_n' stand for to pick the same object in counterfactual situations.

In fact, Geach's mistake is compounded by the fact that, after he acknowledges that 'c_n' refers to Tibbles in the counterfactual situation he considers, he goes on to argue that, since plucking hairs does not generate cats, Tibbles-minus-h_n must already have been a cat in the actual world. This claim trades on another extremely controversial assumption —namely, that proper parts of material objects beloning to a certain kind K that differ in a small respect from the wholes to which they belong are also material objects that belong to the kind K. Van Inwagen (2001) has questioned this assumption arguing that, if it is indeed true, we run into a contradiction.[27]

In order to appreciate this clearly, it is important to bear in mind that, when Geach sets the stage to run his argument, he distinguishes the largest continuous mass of feline tissue on the mat (i.e., Tibbles) from the proper part of Tibbles that contains all of it minus the hair h_n (i.e., Tibbles-minus-h_n). Having assumed that Tibbles and Tibbles-minus-h_n (or, for short, c and c_n) are distinct prior to the plucking of h_n, he then considers what would happen if the hair were plucked. Since he maintains no cats are generated by

[27]The argument that follows is adapted from Van Inwagen. The only difference is that, whereas Van Inwagen considers Descartes and what he calls 'D-minus' (i.e., Descartes except his left leg), my argument bears on the objects that Geach considers in his discussion (i.e., Tibbles and Tibbles-minus-h_n).

events of hair plucking, he then presumably agrees that Tibbles survives the plucking of h_n. Moreover, given that Geach assumes that c_n is, both at the actual world and at the counterfactual world where h_n is plucked, a material object, he presumably accepts also that c_n still continues to exist after being separated from the whole c_n.[28] Finally, since Geach argues that c_n must have already been a cat in the actual world (lest we admit that events of hair plucking spawn new cats into existence) and that it is the same cat as c (lest we admit a multitude of overlapping cats sitting on the mat in addition to Tibbles), Geach woud then agree that, prior to the plucking, c and c_n are identical. But, if that is the case, Geach can be charged with endorsing an inconsistent view since he accepts the four following theses:

(T_1) Tibbles before the hair plucking≠Tibbles-minus-h_n before the hair plucking

(T_2) Tibbles before the hair plucking=Tibbles after the hair plucking

(T_3) Tibbles-minus-h_n before the hair plucking=Tibbles-minus-h_n after the hair plucking

(T_4) Tibbles before the hair plucking=Tibbles-minus-h_n before the hair plucking

[28]If Geach were to reject that c_n (i.e., Tibbles-minus-h_n) continues to exist after the plucking of h_n, he would render meaningless his assumption that c_n is a material object, as Van Inwagen (2001: 80-81) points out when he writes: '(...) if you can cause something to cease to exist by detaching from it (or even by destroying) something that was not one of its parts but simply part of its *environment*, while leaving the arrangement of all of its parts wholly unchanged, if you can do that, then, I maintain, you have not got anything that can properly be called a material object.'

In response to this accusation of inconsistency, Geach could reply that the charge only goes through if we assume that identity is absolute. If we grant, as he does, that there are different relative identity relations, the inconsistency might be resolved by arguing that, though Tibbles before the hair plucking and Tibbles-minus-h_n before the hair plucking are different *masses of feline tissue*, they are one and the same *cat*. However, this reply is unconvincing in light of the fact that Geach acknowledges, as we previously saw, that concepts are mind-made creations that exhibit some vagueness. If that is the case, the relative identity relations that are defined on the basis of concepts such as *cat* or *mass of feline tissue* are bound to be mind-dependent and vague. However, given that Geach appears to be forced to concede the mind-dependency of relative identity relations, we can appreciate that his preferred solution to the puzzle is even more problematic that those he rejects because he forced to admit that there are no objective grounds to assess identity statements: in virtue of his mind creating and deploying a certain concept F, an individual may make the identity statement 'a=b' true whereas another individual may make the same statement false because his mind creates and deploys another concept G. Thus, since the identity statements turns out to be true or false depending on which concept is used and these concepts are mind-made, there is no objective, mind-independent fact that can be used to assess if one of the individuals made a mistake. In light of this, it seems that adopting Geach's view forces us to preclude the possibility of error when making identity (or non-identity) judgments.

I have argued so far that Geach's attempt to show that the sortal relativity of identity is justified insofar as it provides a better solution to the puzzle of

1001 cats than other alternatives is not very convincing. In the remainder of this subsection, I want to show that Geach provides in certain passages of his works certain suggestions to solve the puzzle that do not involve the thesis of sortal relativity of identity.

In one passage that I have quoted earlier where he admits that his theory admits that different A's may be one and the same B, Geach illustrates this by pointing out that 'different intentional objects could be one and the same man'. This passage is interesting because it suggests that Geach acknowledges the possibility of solving the puzzle by claiming that the different masses of feline tissue which are identified as one cat are, in some sense, not material objects like Tibbles but rather some kind of abstract entities. If this is the case, we can give an alternative solution to the puzzle based on this suggestion: we can accept that there is only one material object on the mat that is a cat (namely, c) and accept that this entity has many proper parts c_1, c_2...c_{1000} but deny that these proper parts exist as material objects when they are undetached from c. According to this solution, these undetached proper parts of Tibbles c_1, c_2...c_{1000} are abstract objects that exhibit what Brentano called 'intentional inexistence': they are mental entities that are represent something (namely, Tibbles) but that do not have most of the properties that physical objects such as cats possess (e.g., mass). Since this solution to the puzzle enables us to preserve the commonsense intuition that there is only one cat on the mat and there are reasons to think that the thesis that the proper parts of Tibbles c_1, c_2...c_{1000} are not cats, Geach's use of the puzzle to motivate the sortal relativity thesis is very far from conclusive.

2.4.3 Some traditional replies to RIT_3

Let us consider now the argument for RIT_3. The key premise in this argument is the second one —namely, the thesis that the method of derelativization is the best available option to explain the relation between a sortal term F and the associated relative identity predicate 'x is the same F as y'. According to Geach's view, sortal terms are derivative from primitive relative identity predicates *via* a process in which in relative identity predicates such as 'x is the same F as y' are transformed by derelativization to 'x is the same F as something', which Geach deems to be tantamount to 'x is an F'.

However, a thorough analysis of sortal terms shows that Geach's explanation founders in many cases. In particular, Dummett (1996: 317-319) has presented several counterexamples that throw serious doubts on Geach's argument. For instance, Dummett points out that, even though the predicate 'x is a father' is a sortal insofar as it is associated to a criterion for counting and a criterion of identity, it does not stem by derelativization from the relative identity predicate 'x is the same father as y' but rather from the predicate 'x is a father of y'. In light of this, Dummett goes on to argue that 'x is the same father as y' is not logically prior to 'x is a father' as Geach would suggest; rather, the most plausible way to account for the relative identity predicate 'x is the same father as y' is analyze it as 'x is a father and x is the same man as y'.

In addition to the problem that 'x is a father' raises for Geach, Dummett also observes that there is a second class of counterexamples to the generalization that RIT_3 expresses. This class of counterexamples is illustrated by the predicate 'x is a nationality'. Indeed, even though this predicate is a

sortal insofar as it admits a plural form and it is associated with a criterion for counting and a criterion of identity, one cannot satisfactorily explain how 'x is a nationality' derives *via* derelativization from 'x is the same nationality as y' given that, as Dummett accurately stresses, 'x is nationality' can be better accounted for as being derived from 'x is *of* the same nationality as *someone*' rather than from 'x is the same nationality as *something*'. After pointing out that the pattern of explanation proposed by Geach fails also in this case, Dummett goes on to suggest a different account of how the predicate 'x is a nationality' arises that is inspired by Frege's account of how 'is a direction' may be obtained from 'x is parallel to y':

> We first introduce 'has the same nationality as' or 'is of the same nationality as' as expressing an equivalence relation between human beings, in this particular case one definable as 'is a citizen of the same country as'. Next, we form the operator 'the nationality of', so used that 'the nationality of x is the same as the nationality of y' is equivalent to 'x has the same nationality as y'. Finally, we explain 'is a nationality' as 'is the nationality of someone'. (1996: 319)

As we can appreciate, the account proposed by Dummett, which intuitively explains much better than Geach's proposal how we might obtain expressions such as 'is a nationality', 'is a synonym' or 'is a species' (which correspond to what Dummett calls *abstract count nouns*), does not postulate primitive relative identity relations as the sources of the expressions but rather certain equivalence relations that hold between entities different from the ones to which the expressions in question apply (e.g., persons, meanings

and organisms). In light of this, since we have a plausible account of how many sortal terms arise that does not rely on Geach's proposal, it is clear that Geach's argument fails to establish RIT_3.

2.5 Some novel arguments in favor or RIT

In the previous sections, I have rehearsed the main traditional arguments that support the different components of Geach's RIT and presented some well-known objections to them, intersped with some criticisms of my own. However, I have not dealt with the strongest argument that may be presented to defend the most important component of RIT, which is RIT_1. This argument conists in granting that Quine's method effectively allows us to express identity in theories with finite lexicons, but also emphasizing that, since Quine himself admits that the method fails in theories with infinite lexicons, identity cannot be expressed in these theories since doing this requires us to deploy unrestricted quantification, and this opens the door (as Geach claims) to semantic paradoxes. This shortcoming is particularly problematic for several philosophers to the extent that there seems to be a close connection between absolute identity and unrestricted quantification, as the following passage clearly shows:

> Our sense of a single identity relation that can serve as the target of philosophical discourse is tied to our sense of being able to deploy utterly unrestricted quantification. (...) it seems that our visceral sense that we understand exactly what we mean by 'identity' seems, on the face of it, jeopardized somewhat by those

> philosophical positions that deny the possibility of utterly unrestricted quantification. (Hawthorne 2006: 6fn18)

The intimate connection between identity and quantification is precisely what Geach trades on to set up the first horn of the Dilemma Argument: in order to make a solid case for maintaining that '=' expresses real identity in the axiom schema ID (and not mere indiscernibility vis-à-vis the expressive resources of some theory), we need to be able to quantify unrestrictedly, but we cannot do so without falling prey to semantic paradoxes.

In order to appreciate clearly why the deployment of unrestricted quantification yields a paradox according to Geach, let us assume that the quantifier in the expression 'whatever is true of a' is unrestricted. If that is the case, then it ranges presumably over an all encompassing domain of predicates. Now, if we define F as the predicate 'is identical to something such that nothing that is true of it is also true of a' and we raise the question of whether F is true of a or not, it is not difficult to appreciate that we run run into a situation akin to a semantic paradox: on one hand, the predicate F is true of a iff it is not true of a and, on the other hand, the predicate F is not true of a iff it is true of a.

In order to avoid this type of situation, Geach claims that we have to reject the use of unrestricted quantification —a move that forces us to give up the conceptual resources needed to ensure that we can properly define the notion of identity in theories with non-finite lexicons. Having presented clearly Geach's argumentative strategy, I want to highlight a serious problem that it faces. Geach maintains that unrestricted quantification over predicates cannot be deployed within the context of Quine's method for characteriz-

ing identity under pain of being faced with semantic paradoxes. In light of this, appealing to unrestricted quantification is illegitimate. However, if we cannot legitimately consider an all-encompassing domain of predicates that would be an extension, how can Geach even raise an objection to the use of Quine's method in theories with infinite lexcicons? Indeed, in order to this, he needs to be able to mention this all-encompassing domain, but if he allows himself the expressive resources required to formulate the problem, those same expressive resources can help him provide a solution for it as Hawthorne remarks:

> Geach argues that it is incoherent to quantify over any extension of English in this way [i.e., unrestrictedly]. But didn't Geach have to quantify over extensions of English in order to raise the worry in the first place? Either talk of extensions of English is incoherent, in which case a worry that 'is identical to' doesn't express absolute identity cannot be raised, or else we can raise quantify over a domain of extensions of English, relative to which we can point out that perhaps an I-predicate of English will not express identity proper. But in so far as one can coherently quantify over a domain of extensions, one can stipulatively introduce a predicate that will be immune to the relevant worry: with such quantificational apparatus in place, one can introduce a predicate 'is identical to' stipulating that it is an I-predicate relative to any extension. The apparatus required for raising the worry is the very apparatus needed for solving it. (Hawthorne 2006: 23)

As we can appreciate, the challenge that Geach addresses to advocates of absolute identity can be met by showing that, in order to criticize the use of Quine's method in theories with infinite lexicons by pointing out the dire consequences of unrestricted quantification, he is required to talk about (and, hence, quantify over) an all-encompassing domain. Now, though this response to Geach is certainly effective given the way that he formulates his challenge, Hawthorne observes that there is a way to reformulate Geach's challenge in terms that do not force him to be committed to an all-encompassing domain. Here is how the refomulated challenge goes:

> The most promising version of Geach's objection will allow that there are larger and larger domains of properties available for properties variables, but no maximal domain (...). For any I-predicate introduced by appeal to one domain of properties D_1, one would then always be able to cite a larger domain D_2 relative to which it is intelligible that a pair of objects satisfy the original I-predicate but nevertheless differ with respect to certain properties in D_2. (Hawthorne 2006: 23fn58)

I want to present here a response to the reformulated challenge on behalf of advocates of absolute identity. Following Hawthorne, let us consider a certain theory T_1 that is defined over a domain of predicates D_1 and let us define within T_1 a certain predicate I_1 that satisfies ID for all the properties of D_1 (i.e., an I-predicate in T_1). Let us now consider a theory T_2 that is larger than T_1 insofar as it contains, in addition to T_1, a single mondadic predicate that is not part of T_1. If Geach's reformulated challenge is to have any bite, it must be plausible that some pair of objects in D_2 satisfies the predicate I_1

but nevertheless differ with respect to some property in D_2. However, this claim presupposes that, after moving from T_1 to T_2, the predicate I_1 in T_2 still exists and remains the same entity that it is in T_1. I want to argue that Geach is not entitled to make this assumption, on the basis that it is at odds with what he requires to generate the doubt that I_1 may not express identity in T_2.

Here is my argument. If the meaning of an expression is relative to the theory the expression belongs to, any modification in the theory (such as the addition of new predicates) potentially implies a change of meaning of the expression. This is precisely what Geach relies on to suggest that an I-predicate in T_1 might no longer be an I-predicate in a richer theory T_2. But just as an expression that is an I-predicate in T_1 might cease to be that in T_2, the meaning of any expression that belongs to T_1 (not just the I-predicate) might also change when the expression is considered in T_2.[29]

This relativity of meaning suggests something more radical. Indeed, one may wonder if the expression originally considered in T_1 still exists in T_2, or if what we consider in T_2 is a different expression similar to the one considered in T_1, but nonetheless distinct to the extent that Geach not only takes the meaning of the original expression to be relative to T_1 (thus making possible that the same words employed in the same order have a different meaning in a richer theory), but that he also acknowledges that the fact that the same symbols that make up an expression in T_1 are also used in T_2 in the same order does not guarantee that we are also considering the same expression in

[29]Nelson (1970: 248) appreciates this point when he writes: 'But if the move from T_1 to T_2' does bring with it a change in ontologies, it is also reasonable to think that some of the predicale of T_1, which all occur in T_2, will have different senses, different dictionary readings, in T_2.'

T_2.

In light of this, Geach faces a dilemma. If the meanings of the expressions of a theory T_1 are relative to it, and the occurrence of the same symbols in the same order in an extension T_2 of T_1 does not guarantee the occurrence of the same expression, Geach cannot guarantee that any expression whatsoever in T_1 exists also in any of its extensions, in which case he cannot appeal to the idea that I_1 really denotes indiscernibility rather than identity. But, if the meanings of the expressions of T_1 are not relative to that theory, and one is free to stipulate that the expressions that exist in T_1 exist also in a consistent extension that differs and have the same meanings that they have in T_1, the worry that there may be no way to guarantee that an I-predicate expresses identity rather than mere indiscernibility in T_2 evaporates. Consequently, it is clear that even the reformulated version of Geach's argument fails to establish RIT_1.

2.6 Conclusion

It is time to recap. In this chapter, I have presented the main components of Geach's Relative Identity Theory and rehearsed the traditional arguments that have been used to support it. I have also presented some traditional responses that these arguments have elicited and I have reinforced these responses with some objections of my own to Geach's arguments. From my discussion of Geach's arguments, what clearly transpires is that, even if the some of initial intuitions that motivate Geach's relative identity theory may seem attractive, the main components of the theory (in particular, RIT_1,

which is the most important thesis) are very poorly supported. In particular, even if we reformulate Geach's arguments to try to avoid some possible ojections, it turns out that the reformulated versions of Geach's arguments fail to establish RIT_1. In light of this, it is not surprising that Geach's Relative Identity Theory is almost universally shunned.

Chapter 3

Vindicating the necessity of identity

3.1 Introduction

One of the most important formal results that Marcus (1947) has established is the necessity of identity.[1] This result has garnered enormous admiration from philosophers and logicians in light of the paucity of the assumptions required for it: only two self-evident principles (namely, Leibniz's Law and the necessity of self-identity) are needed to prove it, as the following derivation —which is different from Marcus' proof— shows:

[1]The importance of Marcus' result mainly resides in the fact that, if it is indeed true, it both opens certain avenues of philosophical inquiry and acts as a constraint on the development of others. For instance, it sanctions the cogency of the notion of *a posteriori* necessity, which has been the object of a large body of literature —for further discussion on this notion, see Soames (2006). In addition, it acts as a constraint on the modal status of certain views. For instance, if the thesis that events are identical to property examplifications by substances at a time, which is defended by Kim (1993), turns out to be true, then the thesis is necessary.

1. $(x)\Box(x{=}x)$ Necessity of self-identity

2. $(x)(y)[x{=}y{\rightarrow}(\phi x{\rightarrow}\phi y)]$ Indiscernibility of the Identical

3. $(x)(y)[x{=}y{\rightarrow}(\Box(x = x){\rightarrow}\Box(x = y))]$ 2, Substitution Instance

4. $(x)(y)[x{=}y{\rightarrow}\Box(x = y)]$ 1, 3 since $\Box(x = x)$ is true

Despite the simplicity of the proof, many have disputed the result in different ways. For instance, some have argued that the argument is valid only if (a) we can substitute *salva veritate* co-referential terms within modal contexts and (b) we can make sense of necessary attributions to objects (i.e., attributions such as the property *being necessarily identical to* x). And, since (a) and (b) respectively presuppose the legitimacy of quantification in modal contexts and the cogency of the notion of *de re* modality, Quine's (1943, 1947, 1953) arguments against quantified modal logic have been used to challenge the abovementioned derivation.

In response to Quine's assault on quantified modal logic, Kripke makes two key moves. He first (1980: 34) highlights that many philosophers fail to distinguish the concepts of *analyticity*, *necessity*, *a prioricity* and *certainty*.[2] Having done that, he (1980: 35) suggests that, once these concepts are properly disentangled (in particular, once the concept of necessity is separated from that of analyticity), it is easy to appreciate that not all necessity is *lin-*

[2]Ayer provides us with a good illustration of this view insofar as he maintains that analytic truths are both a priori and necessary. Indeed, after he (1946: 78) defines an analytic proposition as one whose 'validity depends solely on the definitions of the symbols it contains', he argues that, since the truth of analytic propositistions is independent from the nature of the world (which is ultimately contingent), they are necessary. And he also argues that, because the definitions of our symbols are given by linguistic conventions and because we know what our conventions are since we are the ones who establish them, we know the truth of analytic propositions a priori.

guistic. In addition to linguistic or *de dicto* necessity, there is also a notion of *metaphysical* or *de re* necessity (i.e., of necessity that arises from the way things are, not from the meanings of our concepts).

If we embrace the metaphysical notion of necessity, Quine's original worry about the interpretation of modal matrices does not take off the ground because the problem he raises about the legitimacy of quantification in modal contexts only arises within the framework of the linguistic conception of necessity.[3] However, this move does not solve Quine's challenge to the intelligibility of essentialism, which is defined as the view that objects have some of their properties necessarily, regardless of how we characterize them.

Kripke's solution (1980: 48) to this problem consists in arguing that certain terms do reveal more about the essence of the objects they refer to than others because they are *rigid designators* —a rigid designator being a term that refers to the same thing at all the counterfactual situations where that thing exists and does not refer to anything else at counterfactual

[3]This can be clearly appreciated if we bear in mind that Quine worries that there is no proper notion of satisfaction for modal matrices of the form '$\Box\phi x$' on the account that there are failures of substitution of co-referential terms in modal contexts. Now, these failures are taken as evidence of the incoherence of a notion of necessary satisfaction because Quine conceives the notion of satisfaction in modal contexts in 'substitutional' terms: an object o satisfies a modal matrix of the form '$\Box\phi x$' iff there is some name 'a' of o which can be substituted for the 'x' in the matrix and makes the resulting sentence '$\Box\phi a$' true. This is precisely why Quine views cases in which two expressions '$\Box\phi a$' and '$\Box\phi b$' differ in truth value despite the fact that 'a=b' is true as revealing the incoherence of the notion of satisfaction in modal contexts. The way to undercut Quine's argument is, as Stanley (1997: 561) mentions, to reject the 'substitutional' construal of satisfaction in modal contexts and adopt an 'objectual' one in which an object o satisfies a modal matrix of the form '$\Box\phi x$' iff "the object which that assignment assigns to 'x' is necessarily ϕ, that is, is ϕ with respect to every possible situation, *irrespective of any names of the object*." As we can easily appreciate, the move from the 'substitutional' to the 'objectual' interpretation of necessary satisfaction involves switiching from a linguistic conception of necessity to a metaphysical one.

situations where that thing does not exist. Now, if some terms are indeed rigid designators, Quine's second worry may be explained away because rigid designation provides us with the machinery needed to show that essentialism is not an unintellegible view.[4]

Besides defending quantified modal logic (and, along with it, the necessity of identity) by making the two aformentioned moves and introducing the notion of rigid designation, Kripke (1980: 4) defends as well the necessity of identity by drawing a distinction between two theses that are often conflated: (i) the thesis that identical objects are necessarily identical and (ii) the thesis that true identity statements between rigid designators are necessary. This conflation is illustrated by the fact that, even though identity is necessary (as the derivation from 1 to 4 shows), some authors maintain that there are contingent identities because there are identity statements such as 'Benjamin Franklin is the first Postmaster General' that are contingently true.[5]

For Kripke (*ibid.*), the conflation arises because some philosophers do not realize that the existence of contingent identity statements is due to the fact that there are non-rigid designators (indeed, a true identity statement

[4]The basic idea is this: suppose that some term τ is a rigid designator. Since τ designates the same thing O at all the possible worlds where O exists and does not designate anything else at possible worlds where O does not exist, the fact that a sentence of the form '$\Box F\tau$' is true is tantamount to say that O has the property F at all possible worlds where it exists. From this claim, the idea that O has necessarily the property F regardless of how it is described is just a short step away.

[5]Another motivation for the thesis that there are contingent identities springs from the abovementioned failure to distinguish necessity and a prioricity (and the corresponding failure to distinguish contingency and a posteriority). Since certain identity statements such as 'Pain is a brain process' are known only a posteriori, some proponents of mind-brain identity theory such as Smart (1959: 152) claimed, when they asserted the statement 'Pain is a brain process', that the identity asserted was contingent since the statement was established empirically. For further discussion, see Yablo (1987: 294) and Fitch (2004: 87-88).

is contingent because it involves at least one non-rigid designator) and that the existence of non-rigid designators is a contingent fact about natural languages. In virtue of this, the existence of contingent identity statements has no impact on the conclusion of Marcus' argument because (i) is 'a thesis of philosophical logic *independent of natural language*.'

Having shown that one motivation for the view that there are 'contingent identities' stems from the conflation between (i) and (ii), Kripke then distinguishes a further thesis: (iii) the thesis that true identity statements between names are necessary. This thesis follows from (ii) and the central claim that Kripke argues for in *Naming and Necessity*: the thesis that proper names are rigid designators. Using this thesis as well as the parallel thesis that variables are rigid designators,[6] Kripke not only offers a vindication of Marcus' formal result, but he also allows us explain why the identity statments 'Mark Twain is Samuel Clemens' and 'Mark Twain is the author of *Huckleberry Finn*' differ in their modal status (the former being necessarily true while the latter is contingently true) despite the fact that the three singular terms 'Mark Twain', 'Samuel Clemens' and 'the author of *Huckleberry Finn*' refer to the same object.[7]

[6]Even though Kripke does not give an explicit argument for the rigidity of variables that completes his argument for the rigidity of names, such an argument can be reconstructed on the basis of the objectual interpretation of the notion of satisfaction for open modal matrices that Kripke adopts when he accepts that an open modal matrix '$\Box\phi x$' is satisfied iff there is an individual a that the variable x is assigned to in every possible circumstance. For a detailed presentation of the argument, see Stanley (1997: 561-562).

[7]Acccording to Kripke, although the object referred to by both terms in the second statement is necessarily identical with itself, the statement is contingently true because it involves a non-rigid designator (i.e., the definite description 'the author of *Huckleberry Finn*'). Now, it is important to point out that not all definite descriptions are non-rigid. Kripke (1980: 21fn21) remarks that some definite descriptions such as 'the smallest prime' are rigid in virtue of the the fact that the object they denote meets a condition that is satisfied in all counterfactual situations. This type of rigidity, which he labels *de facto*

Although Kripke's defense of the proof of the necessity of identity and of (iii) has been very influential, not everybody has been convinced by his arguments. For instance, Gibbard (1975) argues, *pace* Kripke, that there are certain identity statements involving exclusively proper names that are contingently true. In order to show this, Gibbard presents us with the following thought experiment: he imagines that a clay statue (named 'Goliath') and the lump of clay it is made of (named 'Lumpl') come into existence and cease to exist at the same times. In light of this, there is a strong case to affirm that the statement 'Goliath=Lumpl' is true but, considering that Goliath would have been crushed had it been squashed while Lumpl would have persisted, there is also a strong case to affirm that the statement 'Goliath=Lumpl' is contingently true.

The necessity of identity and (iii) face more recent challenges as well. For instance, Wreen (1998) presents an argument to show that the statement 'Hesperus=Phosphorus' is contingently true, the basic structure of which is the following: given that it is a contingent fact that 'Hesperus' refers to Venus and that it is also contingent that 'Phosphorus' refers to Venus, both names contingently refer to the same object. Now, if 'Hesperus' and 'Phosphorus' refer to the same object, then the identity statement 'Hesperus=Phosphorus' is true. But, since it is contingent that 'Hesperus' and 'Phosphorus' refer to the same object and the claim that 'Hesperus' and 'Phosphorus' refer to the same object is logically equivalent to the claim that 'Hesperus=Phosphorus' is true, then 'Hesperus=Phosphorus' is contingently true.

Following a different line of criticism, Della Rocca (2002) argues that the

rigidity, is distinguished from the *de jure* rigidity of names and variables, which arises from a linguistic stipulation.

proof of the necessity of identity, which crucially depends on essentialism, is subject to self-undermining circularity. To be more specific, Della Rocca claims that Kripke's endorsement of essentialism is problematic because it is concomitant with a rejection of similarity relations to account for *de re* modal properties such as the one expressed by the modal statement 'Humphrey might have won'. And, since Kripke himself employs similarity relations to explain why some modal statements (e.g., the statement 'Hesperus might not have Phosphorus') appear to be true but cannot justify using similarity relations in the latter case and dismissing them in the former without appealing to the necessity of identity, Della Rocca then argues that essentialism and the necessity of identity are locked in an explanatory circle.

The main purpose of this chapter is to vindicate the necessity of identity and (iii) against the abovementioned objections. Thus, I consider in §3.2 Gibbard's Goliath/Lumpl thought experiment in more detail and, after highlighting the reasons that he gives to defend it and to challenge Kripke's view, I show why they are flawed. I examine in §3.3 more in detail Wreen's argument for the claim that 'Hesperus is Phosphorus' is contingent and I show how one may respond to his argument. Later on, I turn in §3.4 to consider Della Rocca's claim that the proof of the necessity of identity is undermined by circularity, and I show how Della Rocca's claim can be answered by advocates of the necessity of identity. Finally, I offer in §3.5 a brief conclusion.

3.2 Gibbard's defense of contingent identity

Gibbard's defense of contingent identity involves showing, *pace* Kripke, that there are some identity statements involving exclusively proper names that are contingently true —i.e., that there are certain counterexamples to (iii). Now, in order to have a thorough picture of Gibbard's objection to (iii), it is imperative to have a good understanding of the argument that Kripke employs to derive it. Considering that this argument crucially depends on the thesis that names are rigid designators, I present in §3.2.1 the main argument that Kripke develops to argue for that thesis. Having done that, I provide in §3.2.2 a detailed account of the Goliath-Lumpl scenario, which is accompanied by a theory of names that is different from Kripke's, and I present the main arguments that Gibbard puts forward to adopt it and to reject Kripke's theory. Finally, I respond to Gibbard's arguments in §3.2.3 and I also argue that the very idea of contingent identity as Gibbard defends it is inconsistent.

3.2.1 Kripke's argument for the rigidity of names

As I mentioned in the introduction, one of the questions that Kripke aims to answer is why two identity statements such as 'Mark Twain is the author of *Huckleberry Finn*' and 'Mark Twain is Samuel Clemens' differ in their modal status (the former being contingent while the latter is necessary) even if all the singular terms involved in them refer to the same object. His solution to this consists in arguing that, while the former identity statement is contingent because it involves the non-rigid designator 'the author of *Huckleberry Finn*',

the latter is necessary because identity statements involving exclusively rigid designators are necessarily true if true at all and because names are rigid designators. Given that the first thesis is rather uncontroversial once we accept the notion of rigid designator, Kripke devotes a good part of his efforts in the first two lectures to defend the thesis that names are rigid designators.

In light of this, I want to consider in detail the argument that Kripke develops for the thesis that names are rigid designators. The argument, which is presented informally by Kripke, is based on the development and the use of a linguistic test aimed at determining if any arbitrary singular term is a rigid designator or not. Given that a singular term τ is a rigid designator iff whatever object that τ refers to in the actual world is also referred to by τ in all the possible worlds in which it exists and τ does not refer to anything other than that object in possible worlds in which it does not exist, the following lingusitic test for rigidity emerges from the previous definition:

> (RT) Any singular term 'τ' is a rigid designator iff the sentence 'τ could not have existed without being τ, and nothing other than τ could have been τ' is true.

After introducing the linguistic test, Kripke (1980: 48) applies it to names such as 'Nixon' and argues that names are rigid because 'although someone other than the US president in 1970 might have been the president in 1970 (e.g., Humphrey might have), no one other than Nixon might have been Nixon.' As we can appreciate, Kripke's test appeals to our semantic intuitions concerning the truth values of the different substitution instances that are

obtained if we replace τ either by a name or by a description in the sentence schema contained in the right-hand side of RT.[8] In the case of names such as 'Nixon', the substitution instances are true according to our intuitions whereas, in the case of definite descriptions such as 'the US President in 1970', the substitution instances are false.[9] Thus, the master argument that Kripke relies on to argue for the rigidity of names can be reconstructed as follows:

> (R_1) Given the definition of a rigid designator, there is a test that allows us to determine if a singular term τ is a rigid designator, and one of its formulations is RT.
>
> (R_2) Our semantic intuitions are such that, when we apply RT to names such as 'Nixon', they come out as being rigid designators.
>
> (R_3) Thus, names are rigid designators.

Now, given that Kripke's argument depends on RT, which in turn is justified on the basis of certain semantic intuitions, the argument has been objected by questioning the correctness of these intuitions. In the following subsection, I present Gibbard's criticism of Kripke's argument, which involves

[8]Reliance on semantic intutions has been criticized by some as a doubtful methodology, particularly because intuitions are traditionally viewed in a Cartesian framework as the outcome of some mysterious faculty of the human mind. But there are naturalistic accounts of intutions. For a recent example that treats semantic intuitions as a certain kind of empirical central-processor responses, see Devitt (2006).

[9]Kripke claims that names are rigid even if they contain descriptive elements (e.g., 'Jack the Ripper' or 'Attila the Hun'). In order to explain this, he (1980: 79-80) suggests that the descriptive elements only help to fix the reference of the name in the actual world, but they do not give its meaning. Thus, the sentence 'Attila the Hun could not have existed without being Attila the Hun, and nothing other than Attila the Hun could have been Attila the Hun' is true even if the man we call Attila was not actually a Hun.

challenging these intuitions on the basis of a hypothetical scenario in which a statue (called 'Goliath') and a lump of clay (called 'Lumpl') are contingently identical and developing of alternative theory of names according to which names are rigid with respect to sortal concepts.

3.2.2 The Goliath-Lumpl scenario and Gibbard's theory of names

3.2.2.1 The Goliath-Lumpl scenario

Gibbard initially introduces his case for contingent identity by claiming that, under certain circumstances, a clay statue is identical with the piece of clay it is made of. And he further adds that, if the statue is identical with the piece of clay, the identity is contingent. Thus, granting that the statue is s and the piece of clay is c, the thesis that Gibbard seeks to establish can be expressed as follows:

(CI) $s{=}c$ & $\Diamond(s$ exists & c exists & $s{\neq}c)$

After formulating clearly the thesis, Gibbard remarks that, if true, CI has several important ramifications. Indeed, not only does it raise a counterexample to Kripke's defense of Marcus' result, but it also suggests an alternative theory of names which has, according to Gibbard, many advantages with respect to that of Kripke. Thus, even though Gibbard acknowledges that Kripke has effectively shown that many of the traditional motivations to endorse the view that there are contingent identities (e.g., the conflation of the distinction between a priori and a posteriori truths with the distinction

between necessary and contingent truths or the descriptive theory of names) are mistaken, he also argues that, *pace* Kripke, there are some contingent identity statements that involve exclusively names.

Here is how Gibbard proceeds. After pointing out that CI is to be interpreted as stating a claim about strict, timeless identity (and not about mere similarity or some other relation),[10] he asks under which circumstances the clay statue s is identical to the piece of clay c. Since a necessary condition for c and s to be identical is that c and s begin to exist and cease to exist at the same times, Gibbard claims that we need to provide *persistence criteria* for both pieces of clay and clay statues.[11]

[10]In light of this, it is quite puzzling to see that, immediately after that reminder, Gibbard (1975: 188) makes the following statement: 'For two things to be strictly identical they must have all properties in common.' Indeed, if CI is a claim about strict identity then it cannot involve two things since strict identity between s and c amounts to the fact that s and c are one and the same thing. As I argue in more detail further down (see the Reply to the Cross-World Identities Worry argument in §3.2.3.1), I take this assertion to be not a mere oversight from Gibbard, but rather an indication of the fact that the very idea of contingent identity is inconsistent.

[11]The strategy that Gibbard deploys (i.e., using persistence criteria in order to answer a question about identity) is problematic because answering identity questions such as 'Under what circumstances are s and c identical?' by appealing to some persistence criteria (which are specific to kinds of entities) requires, as Johnston (1987: 123) remarks, substituting questions about identity by questions about kind membership such as 'Under what circumstances does an object c of the kind $\mathcal{C}$ (with persistence criteria so-and-so) also belong to the kind $\mathcal{S}$ (with persistence criteria thus-and-thus)?' The main difficulty with this move is that the main motivation behind it is a certain reductionist picture according to which, using Oderberg's (1993, 3) words, 'metaphysical analysis reveals to us that (...) some other condition obtains in virtue of which identity holds in any particular case' and this reductionist picture is fraught with problems. One major problem, which is discussed by Kripke in his 1978 unpublished Cornell lectures on time and identity, is that if we provide a reductive identity criterion for a certain kind of entities (say, instants or times) in terms of a second kind of entities (say, events), we are bound to provide a reductive identity criterion for the second kind of entities presumably by appealing to a third kind of entities, and so on. However, this cannot go indefinitely under pain of violating the axiom of well-foundedness for sets. In light of this, Kripke concludes that, if we subscribe to the abovementioned reductive strategy, we are forced to acknowledge that there is some kind of entities for which identity is primitive.

In the case of pieces of clay, Gibbard suggests that a piece of clay persists as longs at its constituent parts are bound by a certain *sticking* relation whereas a clay statue persists as long as the piece of clay is made of persists and maintains the same shape that it has.[12] Even though these criteria seem to fit roughly our discourse regarding pieces of clay and clay statues, Gibbard is well aware that some may find them too strict and, in order to block a potential objection, he claims that his argument for CI does not depend on the specific criteria he mentions in any crucial way: since the criteria can be taken to be purely stipulative, they can be replaced by other alternative criteria that also correspond to the way we talk about pieces of clay and clay statues. Thus, the selection of persistence criteria is arbitrary to a certain extent. The only constraint that Gibbard imposes on the selection is that, whatever the criteria are, they must be consistent with the claim that both pieces of clay and clay statues are objects and that both can be designated by names.

Now, if pieces of clay and clay statues are considered along the lines of the aformentioned presistence criteria, Gibbard (1975: 190) observes that they are 'often distinct' since a piece of clay typically originates as it is broken off from a bigger piece of clay and a statue typically comes into being after the piece of clay is it made of is shaped in a particular way and gets its

[12]The sticking relation that holds among the parts of a piece of clay according to Gibbard is similar to the unity relation that holds between the temporal stages of person according to Perry (1972). In fact, each one of these relations may be considered as a distinct supervenience base for the identity relation —a supervenience base that is specific to a certain kind of objects. But, even if we grant that identity does admit different supervenience bases which are relative to the different kinds of objects to which identity applies (a thesis that I am not inclined to agree with), it does not follow that identity itself is relative to the supervenience bases. For further discussion on this, see Sider (1999: 922-925) and McGinn (2000: 6).

finishing touches.[13] Indeed, they not only begin to exist at different times, but clay statues depend on the preservation of their shape for their continued existence whereas lumps of clay do not.

Although pieces of clay and clay statues are distinct in virtue of the aforementioned features, Gibbard wonders if there are circumstances in which they simultaneously start to exist and in which they also simultaneously cease to exist. If that were so, those circumstances would presumably provide strong support for defending the first conjunct of CI. In order to illustrate one such circumstance, Gibbard (1975: 191) presents us with a hypothetical case in which he makes a clay statue of the infant Goliath by first crafting its two halves and, once they are finished, sticking them together —an act which brings into existence a new piece of clay and the statue. Subsequently, he smashes his creation, which brings and end to both the statue and the piece of clay. In light of this, the clay statue statue and the lump of clay have the same temporal extent.

In the context of this hypothetical scenario, Gibbard affirms that it is quite plausible to affirm that the statue (which is referred to as 'Goliath') is identical to the piece of clay (which is called 'Lumpl') because they not only begin to exist and cease to exist at the same times, but they also have the same mass, shape, location and color at every moment of their mutual history. Of course, he also acknowledges that other explanations of these

[13]It is interesting to notice here that Gibbard assumes that the distinctness relation (which presumably obtains between any x and any y when identity between x and y does not obtain) clearly involves a temporal element. This appears to be at odds with the fact that he takes the identity sign in CI to be read as expressing strict, timeless identity. Now, if identity is to be taken in a strict, timeless sense, distintinctness should also be taken in a strict, timeless sense. As I mentioned above in footnote 10, I take these tensions in Gibbard's claims to be a symptom of the inconsistency of the view he is arguing for.

facts are available. In particular, Goliath may be said to be *constituted* by Lumpl,[14] but this explanation has the inconvenient of creating an ontological inflation given that accepting the claim that Goliath is constituted by Lumpl requires us to assume that Goliath and Lumpl are coincident but nevertheless distinct entities.[15] Considering this, there is some initial plausibility (on the grounds of ontological parsimony) for the hypothesis that the identity sentence 'Goliath=Lumpl' is true.

In addition to the fact that Goliath and Lumpl share all of their non-modal properties, Gibbard mentions that there other reasons that support the thesis that 'Goliath=Lumpl' is true. The first one is that the thesis fits a systematic account of clay statues and pieces of clay in the sense that, since clay statues are tyically molded from a pre-existing piece of clay, clay statues are generally temporal segments of pieces of clay. According to this account, even though in most cases clay statues are preceded by the pieces of clay they are made of, in certain cases (such as the hypothetical scenario that Gibbard considers) the temporal career of a statue coincides exactly with that of the piece of clay it is made of —a fact which strongly suggests according to Gibbard that they are identical.[16]

[14]For a detailed presentation and defense of the relation of constitution, see Baker (1999) and (2002), Wiggins (2001) and Johnston (2005).

[15]Although Baker, Johnston and Wiggins agree that constitution implies distinctness, Noonan (1993) has made an interesting attempt to account for the fact that Goliath and Lumpl begin to exist at the same time and that they have the same shape, location, color and so forth throughout their history in terms of the relation of constitution without paying the ontological cost by arguing that constitution *is* identity. However, there are strong reasons to maintain that Noonan's thesis is erroneous. In particular, while identity is reflexive and symmetric, constitution is irreflexive and asymmetric.

[16]An important point to notice here is that, although Gibbard attempts to provide support for the thesis that Goliath and Lumpl are identical on the basis of (a) the exact spatiotemporal coincidence of Goliath and Lumpl throughout their existence and (b) a metaphysical view that acknowledges the existence of temporal segments or parts of con-

The second reason is that the thesis coheres with the best hypotheses that we have concerning the nature of concrete things as being part of the physical world. Indeed, considering that the most promising accounts of concrete things postulate the existence of some kind of fundamental physical entities of which they are made, Gibbard remarks that, if the thesis is correct, then it should be consistent with these accounts. And that seems to be the case, regardless of whether we take the fundamental physical entities to be point-instants or particles (which are two of the most plausible hypotheses). For Gibbard, if a concrete thing is just a particular set of point-instants,[17] the thesis that Goliath is identical to Lumpl is consistent with the hypothesis that the statue and the piece of clay are the same set of point instants and, if a concrete thing is just a set of changing particles, the thesis that Goliath is identical to Lumpl is consistent with the hypothesis that the statue and the piece of clay are made of the same particles at each instant.

Having provided further evidence for the claim that 'Goliath=Lumpl' is true, Gibbard (1975: 191) then argues that, even if the identity sentence is true, there are strong reasons to maintain that it is contingently true

crete objects, accepting both (a) and (b) is perfectly consistent with other theories that construe the relation between Gibbard and Lumpl as being one of distinctness. In particular, accepting both (a) and (b) is consistent with theories according to which Goliath is constituted by Lumpl. It is only when further assumptions are made (in particular, when one embraces the idea that an object is nothing over and above its parts) that (a) and (b) provide substantial support to the identity thesis. For further discussion on this, see Hawthorne (2006: 88-90).

[17]This claim sounds at first sight paradoxical to the extent that sets are traditionally deemed to be abstract objects. Given the lack of textual evidence, it is difficult to determine whether this claim is just a mere oversight or whether Gibbard really believes that sets are concrete, spatiotemporal entities, as Maddy (1990) does. Notice that, if the latter alternative is correct, Gibbard would be required to provide answers to questions about the locations of different sets, which is clearly an additional explanatory burden on his view.

because Goliath and Lumpl have different *modal* properties. Indeed, if both Goliath and Lumpl had been created along the aforementioned lines but, instead of letting the clay dry, Gibbard had squeezed the clay into a ball then, according to the persistence criteria previously given, Goliath would have ceased to exist whereas Lumpl would have persisted under a new shape.

If this counterfactual scenario is consistent, Gibbard can indeed claim to have established a particular instance of CI in which '*s*' is replaced by 'Goliath' and '*c*' by 'Lumpl'. Now, as I mentioned earlier, Gibbard remarks that the thesis that 'Goliath=Lumpl' is true, albeit contingently, is at odds with Kripke's view of names as rigid designators. Consequently, since the two views conflict, Gibbard presents an alternative theory of names which aims to explain the contingent status of the identity statement 'Goliath=Lumpl'.

3.2.2.2 Gibbard's theory of names

In order to provide motivation for his alternative theory, Gibbard considers Kripke's claim that, if names such as 'Goliath' or 'Lumpl' are rigid designators, any of them refers to the same thing that it refers to in the actual world in all the possible worlds in which that thing exists. But what does it mean for 'Goliath' to refer in all the possible worlds in which it exists to *the same thing* that it refers to in the actual world? As Gibbard remarks, both 'Goliath' and 'Lumpl' refer in the actual world to the thing that he made and then broke. But what would that thing be in the counterfactual situation in which the clay was squeezed into a ball before it had the chance to dry? Gibbard (1975: 194) suggests here that, since asking what a thing would be in a counterfactual situation requires referring to it as a thing of

a certain kind (e.g., as a statue or as lump of clay), asking about the counterfactual goings-on of a thing independently of the way it is referred simply makes no sense.

In light of this, it is clear thar Gibbard endorses the same thesis that underlies Quine's rejection of essentialism: objects do not have any modal properties independently of the way they are referred to. But Gibbard goes beyond the original Quinean anti-essentialist thesis as he develops a theory of names that is at odds with Kripke's. Indeed, while Kripke favors a Millian view according to which the only semantic contribution that a name makes to a sentence in which it occurs is its referent,[18] Gibbard argues that names also have other functions besides denoting their referents (if there have any). The key elements of his proposal are introduced in the following passage:

> Proper names like 'Goliath' or 'Lumpl' refer to a thing as a thing of a certain kind: 'Goliath' refers to something as a statue; 'Lumpl' as a lump. For each such kind of thing, there is a set of persistence criteria, like the ones I gave for statues and for lumps. In rare cases, at least, one thing will be of two different kinds, with different persistence criteria, and whereas one name refers to it as a thing of one kind, another proper name will refer to as a

[18] Although Kripke's key role in the rejection of the descriptivist view of names and the concomitant revival of Millianism is widely acknowledged, it is sometimes forgotten that the view that he favors is not *purely* Millian in the sense that he rejects an unrestricted substitutivity principle such that co-referential names would be interchangeable everywhere not merely *salva veritate* but also *salva significatione*. Since the actual view that he favors emphasizes the *consistency* of the theses that names are modally rigid and that co-referential names are intersubstitutable in modal contexts with the thesis that names are not intersubstitutable in epistemic contexts, his position about names is an intermediate one between Frege and Mill. For further discussion, see Kripke (1979).

> thing of a different kind. In such cases, the identity formed with those names is contingently true. (1975: 194-195)

As the former passage shows, besides denoting a certain thing (if it denotes anything at all), a name has two further roles according to Gibbard: it specifies a kind or a sortal concept under which the thing it refers to falls and it also invokes some persistence conditions that are associated with the sortal.[19] Both of these features are crucial to explain why some identity statements involving exclusively names are contingently true. Indeed, granting that objects have no modal properties *per se* but only relative to the way they are designated, it is clear that, even if we suppose that two names are co-referential in the actual world, the persistence conditions and the sortal concepts that are respectively associated with these names may pick out distinct objects in counterfactual situations. Thus, it is the distinct persistence conditions and the sortal concepts that are respectively invoked by 'Goliath' and 'Lumpl' that explain why, even if 'Goliath=Lumpl' is true, the identity statement is contingently true.

Besides attributing further roles to names beyong mere denoting, Gibbard's theory of names also provides an account of how their references are

[19]Even if we initially grant to Gibbard that a name invokes certain persistence conditions, doubts remain as to what exactly invokes the specific persistence conditions that are associated with a name such as 'Goliath': is it just the mere typographic occurrence of the name? Or is it the use of the name in a certain context of utterance? The first alternative is implausible considering that 'Goliath' is used to denote, not only a clay statue, but also other things belonging different kinds (e.g., a certain man who was a Philistine warrior). But the second alternative is also problematic in the sense that, if what invokes certain persistence conditions for a name is the use of the name in a certain context of utterance, what invokes the persistence criteria falls within the sphere of *pragmatics* —and thus lies beyond the scope of the semantics that Gibbard offers in the appendix of his paper to accommodate modal talk. In light of these problems, Gibbard's thesis that names are associated with certain persistence criteria should be regarded with some reserve.

fixed in counterfactual worlds as well as in the actual world. In the case of counterfactual worlds, Gibbard (1975: 196) suggests that the reference of a name is the product of two factors: what its reference is in the actual world and the persistence criteria that are invoked by the name. In particular, the reference of a name in the actual world determines how the thing that it refers to in a counterfactual world originates, and the persistence criteria that are invoked by the name allow us to determine which thing (among those that have the same origin) the name refers to. To illustrate this, consider again Gibbard's story: granting that 'Goliath' refers at the actual world $W_@$ to the thing that was made by joining two halves and that was subsequently broken, 'Goliath' then refers in a counterfactual situation W_1 to something that originated in the same way as the thing that it refers to at $W_@$ and the persistence criteria associated to it help us determine which thing (among the many that originate in the same way at W_1) the name refers to.

In the case of the actual world, Gibbard's proposal builds on Kripke's causal view (1980: 91-92) according to which a name refers to a certain object if there is a suitable causal chain that links the person using the name with the thing referred to. For instance, both 'Goliath' and 'Lumpl' refer, according to Kripke, to the thing that was made and broken in the actual world because Gibbard was the one who orginally introduced the names to refer to it, thus grounding the causal chain. Now, what Gibbard adds to this view is that persistence criteria often play a role in originally fixing the reference (i.e., in grounding the causal chain). In particular, Gibbard suggests that 'Goliath' refers in the actual world $W_@$ to the *statue* that was mas made and subsequently broken because, in using the name 'Goliath' to refer to the

object that was made and then broken, certain specific persistence criteria are invoked and it is the statue (and not the piece of clay) that satisfies these persistence criteria.

To sum up, there are at least two main features that characterize Gibbard's theory of names: (1) besides referring to their denotations (if they do refer at all), names invoke certain sortal concepts and persistence criteria and (2) these persistence criteria play a role in fixing the reference of the names not only in the actual world but also in counterfactuals situations. It is precisely because of these two features that Gibbard's theory is incompatible with Kripke's: on one side, the claim that that names invoke certain persistence criteria that are associated to them is just inconsistent with the Millian spirit underpinning Kripke's view and, on the other side, although Kripke might agree that persistence criteria can be used to fix the reference of a name in the actual world, allowing persistence criteria to fix the references of names in counterfactual worlds amounts to give into the demand that the references of names are determined in counterfactual situations through certain descriptions —a demand that Kripke flatly rejects as it is grounded on a misleading view of counterfactual situations.[20]

[20]Since allowing persistence criteria to fix the reference of names in both actual and counterfactual situations boils down to make relations of cross-word identity relative to these criteria, Gibbard's theory of names turns out to be similar to Lewis' proposal (1968). Indeed, according to Lewis, the references of names in counterfactual situations are fixed, not by persistence conditions, but by different counterpart relations —which are similarity relations. Thus, Lewis would agree with Gibbard's claim that, even if 'Goliath' and 'Lumpl' are co-referential in the actual world, they have distinct references in certain counterfactual situations because therein 'Goliath' refers to statue-counterparts whereas 'Lumpl' refers to lump-counterparts. But he would sharply disagree with Gibbard's claim that cross-world identity is relative to a sortal on the account that, for him, individuals are worldbound and that, consequently, there is no cross-world identity. For further discussion, see Lewis (1986), §4.3

Now, considering that there is a widepread endorsement of Kripke's view, Gibbard is well aware of the burden of proving that his theory is better. Accordingly, he presents a series of arguments aimed at bolstering his proposal. These arguments can be roughly divided into three categories: those which aim to establish that his theory allows us to account for the same 'data' as Kripke's (and that, accordingly, there is no loss of explanatory power when switching from Kripke's view to his), those which aim to prove that certain aspects of Kripke's view are problematic and, finally, those which aim to reply to certain potential criticisms that can be addressed to his theory. In the following subsection, I offer a more thorough review of the arguments.

3.2.2.3 Gibbard's arguments

The No Explanatory Loss Argument. This argument aims to show that, just as Gibbard's theory enables us to provide an account of why the identity statement 'Goliath=Lumpl' is contingently true, it also explains why the identity statement 'Hesperus=Phosphorus' is necessarily true. According to Gibbard (1975: 198), this is due to the fact that, whereas 'Goliath' and 'Lumpl' respectively invoke different sets of persistence criteria which select different objects in some counterfactual situations, 'Hesperus' and 'Phosphorus' invoke the same set of persistence criteria (namely, that for heavenly bodies). Thus, since 'Hesperus' and 'Phosphorus' refer to one and the same object in the actual world and both names invoke the same persistence criteria, the reference of 'Hesperus' is the same as that of 'Phosphorus' in all counterfactual situations.[21]

[21]The force of this argument stems from the fact that it allows Gibbard to explain using his theory of names not only why the identity sentence 'Hesperus=Phosphorus' is

The Failed Rigidity Test Argument. This argument attempts to establish that Kripke's argument for the thesis that names are rigid designators is flawed because the test RT does not establish that names are rigid designators. More specifically, Gibbard (1975: 199) argues that, when RT is applied to 'Nixon', all that the test establishes is that Nixon might not have been a *man* distinct from the man he in fact was. Considering this, the test merely shows according to Gibbard that 'Nixon' is rigid with respect to the sortal *man*, but not that it is rigid *simpliciter* as Kripke claims. In order to establish that 'Nixon' is rigid, Gibbard argues that Kripke would have to show that Nixon might not have been a distinct *entity* from the one it actually is, but he claims that this thesis is not properly supported by our intuitions that everyday talk of 'the same person' fits well into systematic talk of 'entities'.

The Cross-World Identity Worries Argument. This argument consists in claiming that Kripke's strategy to settle questions of cross-world identity fails and that, in virtue of this, the claim that names are rigid designators *simpliciter* is not true. The basis of this argument is Kripke's thesis that, if names are indeed rigid designators, the cross-world identity question "How do we identify which thing 'Nixon' refers to (if any) in some counterfactual situation W'?" becomes moot: considering that 'Nixon' just refers in W' to the thing that it refers to in the actual world $W_@$, there is no need to identify which thing it refers to in W'.[22] In fact, to provide further support

necessarily true but also why other identity sentences that Kripke might use as illustrations of (iii) are necessarily true. Indeed, Gibbard may argue that the identity sentence 'Samuel Clemens=Mark Twain' is necessarily true because both names refer in the actual world to the same thing and both invoke the same persistence criteria —namely, persistence criteria for persons.

[22]For a more elaborate formulation and discussion of the problem of cross-world identity,

to the claim that crossworld identity worries are moot, Kripke (1980: 44) provides an account of their source by arguing that they stem from a certain mistaken picture of counterfactual situations. For Kripke, cross-world identity worries arise if we conceive counterfactual situations as entities the contents of which are identified through powerful telescopes. But, if we conceive counterfactual situations as states of affairs which are *stipulated* rather than discovered, cross-world identity worries dissolve.

Now, Gibbard argues by *modus tollens* that, since cross-world identity worries still arise even if we suppose (following Kripke) that counterfactual situations are stipulated, the thesis that names are rigid designators is simply not true. Specifically, he (1975: 200) asserts that the Goliath-Lumpl scenario constitutes a clear counterexample to Kripke's claim that stipulation of counterfactual situations allows us to quell cross-world identity worries. Indeed, according to him, the counterfactual scenario W_1 where Goliath and Lumpl are brought into existence as they are in the actual world $W_@$ but in which the clay is squeezed into a ball before it has the chance to dry is fully stipulated, but it fails to provide a definitive answer to the question 'Which of the two is the one thing that, in the actual world, I made and then broke?'

The Leibniz's Law Argument. This argument is intended to respond to what Gibbard takes to be the most prominent objection to his theory. The key element underpinning this objection is Leibniz's Law, which enables potential critics of Gibbard to derive a result that undermines CI. Indeed, given that Leibniz' Law (or, to be more precise, the Indiscernibility of the Identical) states that if x and y are identical then whatever property ϕ that

see Plantinga (1974), chap. 4, sect. 3.

x has is also had by y, from the following thesis (which Gibbard accepts)

(L_1) $\Box$(Lumpl exists $\rightarrow$ Lumpl=Lumpl)

we can derive, using the relevant instance of the Indiscernibility of the Identical and the assumption that Goliath=Lumpl, the following conclusion

(L_2) $\Box$(Lumpl exists $\rightarrow$ Goliath=Lumpl)

which is at odds with CI. To answer the objection, Gibbard argues that the inference from L_1 to L_2 is valid only if the modal context '$\Box$(Lumpl exists $\rightarrow$ _____=Lumpl)' expresses a genuine property. Now, given that a genuine property of an object *o* is a property that *o* has *independently of the way it is described*, the property that the abovementioned modal context expresses is genuine, according to Gibbard (1975: 201), only if 'it gives something that is true of Lumpl or false from Lumpl independently of the way Lumpl is designated, and whether it does is the point in question.' To sum up, Gibbard argues that the inference from L_1 to L_2 is valid only if one assumes essentialism but, given that assuming esentialism amounts to beg the question against contingent identity, critics cannot rely on the inference to undermine CI.

The No Diminished Expressive Power Argument. The purpose of this argument is to show that adopting Gibard's theory does not require paying too high a price, as some would argue. Considering that the response to the objection based on Leibniz's Law requires Gibbard to rely on Quine's anti-essentialist thesis according to which objects have no modal properties independently of the way they are described, Gibbard is concerned with

avoiding a consequence that may be drawn from this thesis —namely, that expressions such as

(L_x) $\Box$(Lumpl exists $\rightarrow$ x=Lumpl)

which involve one or more free variables within the scope of a modal operator are ill-formed. Gibbard (1975: 202) wants to eschew this consequence insofar as he worries that some may argue on this basis that, if contingent identity is endorsed, 'we shall be unable to say many things we need to say, both in scientific talk and in everyday life.' To answer this challenge, Gibbard develops a semantic theory for modal discourse that is based on the key insight that Carnap (1947) relied on to defend modal logic from Quine's onslaught: having the variables embedded in modal contexts range, not over individuals, but over individual concepts and having the predicates embedded in modal contexts refer to, not sets of objects, but second-order concepts.[23] Using this insight, Gibbard argues that expressions in which variables occur within the scope of modal operators are well-formed; the effect of modal operators is just to 'switch' the universe of discourse over which variables normally range.

To illustrate this, Gibbard asks us to consider the formula '$Ex \rightarrow Hx$' where 'E' stands for *exists* and 'H' stands for *is human-shaped throughout its early history*. If a modal operator (say, '$\Box$') is used on the formula,

[23]In addition to Carnap, Church (1943: 46) presents a very similar proposal, claiming that the conclusion that Quine should have drawn from his logical argument against quantified modal logic is that 'variables must have an intensional range —a range, for instance, composed of attributes rather than classes.' Although the formulation of Church's proposal is problematic because his talk about 'intensional ranges' is at odds with Cartwright's (1967) persuasive arguments that show that intensionality is a feature of contexts rather than ranges, I will not delve into this issue here.

then the variable 'x' does no longer admit objects as substituends but rather ranges over individual concepts such *Goliath* or *Lumpl* and the predicate constants no longer have the abovementioned references, but rather stand for the second-order concepts *is a concept of an individual that exists* and *is a concept of an individual that is human shaped throughout its early history* respectively. Thus, the semantics that Gibbard provides allows him to answer the previous objection in the following way: although he admits that objects do not have modal properties *per se*, he acknowledges that they have modal properties relative to certain sortals, and these properties are the ones that our modal discourse refers to.

In other terms, if I utter the sentence 'Paris could have been smaller than Reims', I attribute to Paris a certain modal property according to Gibbard — not the property 'being smaller than Rheims' but rather the property 'being smaller than Rheims *qua* city'. Granting that 'S' stands for the two-place predicate 'is smaller than', that r stands for 'Rheims' and that '$\mathcal{C}$' stands for the predicate 'is a city-rigid concept', the property that is attributed to Paris by uttering the sentence may be formally expressed as follows:

(P) $(\exists y)(x{=}y \ \& \ \mathcal{C}y \ \& \ \Diamond Syr)$

It is important to notice that, as the variable 'x' in P lies outside the scope of modal operator, it ranges over the standard universe of discourse (namely, individuals) whereas 'y' ranges over individual concepts. In light of this, since both variables range over distinct domains, the identity sign '=' in P does not express identity —an observation that Gibbard (1975: 206) generalizes to all cases in which the identity sign occurs inside a modal

context. Now, considering that the identity sign does not express identity in modal contexts, Gibbard is concerned that some may argue, following Quine, that his reliance on Carnap's proposal forces him to change the subject. To this worry, he (*ibid.*) responds that, even if the names 'a' and 'b' refer to individual concepts when they occur in modal contexts, the identity sentence 'a=b' in a modal context expresses the identity of individuals that 'a' and 'b' refer to because the sentence means 'that a and b are concepts of the same individual [and] that amounts to saying that the individual of which a is the concept is identical to the individual of which b is the concept.'

There is a further concern that Gibbard mentions. One of the consequences of adopting Carnap's proposal consists in embracing a form of essentialism for individual concepts: even though we cannot state that the thing that Gibbard actually made and smashed would have persisted after been squashed (one can only say that of the thing *qua* lump), we can say that the individual concept *Lumpl* necessarily falls under the second-order concept *is a concept of an individual that is a lump*. Since Gibbard is well aware that his rejection of essentialism for objects and his acceptance of essentialism for individual concepts may be seen as yielding an inconsistent view, he (1975: 208) justifies this discriminatory treatment arguing that, as concepts are just functions that map a world to an individual in that world, 'there is no problem of what that function would be in a possible world different from the actual one.'

3.2.3 A refutation of Gibbard's views

In this subsection I have two goals. My first order of business is to provide a systematic refutation of Gibbard's arguments in defense of CI and his theory of names. After doing that, I present an argument that establishes that Gibbard's views are incompatible. On this basis, I conclude that Kripke's defense of Marcus' result and his argument to demonstrate (iii) are not undermined by Gibbard's challenge.

3.2.3.1 Replying to Gibbard's arguments

Reply to the No Explanatory Loss Argument. The gist of the No Explanatory Loss Argument is that, as Gibbard's theory allows us to account for the same 'data' that Kripke's view accounts for, there is no loss of explanatory power if we substitute the latter by the former. Now, this argument depends on the thesis that names invoke certain persistence conditions rather than others. Indeed, Gibbard claims that his theory can account, just as Krikpe's view, for the fact that the identity statement 'Hesperus=Phosphorus' is necessarily true because, in addition to denoting the same individual in the actual world, both 'Hesperus' and 'Phosphorus' invoke the same persistence conditions —namely, those for heavenly bodies.

But assuming that 'Hesperus' and 'Phosphorus' invoke the same persistence conditions is doubtful for at least two reasons. First, one may counter that this assumption is false. Just as Gibbard argues that 'Goliath' and 'Lumpl' invoke distinct persistence conditions, one may argue that 'Phosphorus' and 'Hesperus' also invoke distinct persistence conditions —conditions for *visible-at-dawn heavenly bodies* and conditions for *visible-at-dusk heav-*

enly bodies respectively. If this is indeed the case, then the identity statement 'Hesperus=Phosphorus' fails to be necessarily true by the standards of Gibbard's theory.

Gibbard may reply to the previous objection by arguing that *visible-at-dawn heavenly body* and *visible-at-dusk heavenly body* are gerrymandered kinds that fail to have well-defined persistence conditions for the objects that fall under them. However, this reply is effective only if Gibbard can supply a well-grounded distinction between genuine kinds and gerrymandered ones. Now, in light of the strong anti-essentialist stance that he adopts with respect to objects, it seems that any distinction that he can draw between genuine kinds and gerrymandered ones has to be given in terms of similarity relations. But, since similarity is, as Goodman (1972) has argued, an inherently relative and variable relation, it cannot provide a well-grounded basis to distinguish genuine kinds from gerrymandered ones.[24]

A second reason to doubt that 'Hesperus' and 'Phosphorus' invoke the same persistence conditions is the following. Let us grant that, in addition to referring to the same object (i.e., Venus), 'Hesperus' and 'Phosphorus' do invoke certain persistence conditions and that these persistence conditions are supplied by a certain descriptive content associated with each name. However, this supposition is not enough to support the thesis that 'Hesperus' and 'Phosphorus' invoke the *same* persistence conditions. Here is why. In an

[24]Somebody might interject here the following question on behalf of Gibbard: aren't there any objective, primitive similarities that can be used to ground a distinction between genuine and gerrymandered kinds? Some nominalists have considered seriously this alternative, but pursuing it involves high costs. In particular, primitive similarity relations are very complex and artificial, as Lewis (1983a: 347-348) has shown. In light of this, I do not think that primitive similarity relations are good candidates to ground a distinction between genuine kinds and gerrymandered ones.

effort to accomodate certain descriptivist intuitions from a Kripkean standpoint (in particular, the intuition that the meaning of a sentence involving a proper name is not exhausted by its semantic content), Soames (2002) has suggested that, in sentences involving proper names, it is important to distinguish their semantic contents (which he identifies with Russellian propositions) from certain descriptive contents that the sentences are used to convey when uttered by speakers in certain contexts.

To illustrate this distinction, Soames asks us to consider the sentence 'There is an article about David Lewis in the *New York Times*'. For him, this sentence expresses a Russellian proposition, but it also may be used to convey additional descriptive information. For instance, if it is uttered in the context of a conversation between Princeton faculty members, it may be used convey information regarding the fact that there is an article about the *philosopher* David Lewis in the New York Times, or information about the fact that there is an article in the *New York Times* about their *friend* David Lewis, or information about the fact that there is an article in the *New York Times* about their *colleague* David Lewis, and so on. Now, after introducing the distinction, Soames (2002: 82) uses it to argue that, when one asserts a sentence involving a proper name, the sentence not only expresses a determinate Russellian proposition: it also conveys a certain descriptive content but 'in many cases there seems to be no uniquely correct way of specifying precisely what (extra) descriptive content gets into one's assertion.'

If Soames's proposal regarding the indeterminacy of the descriptive content conveyed by sentences involving names is correct, then we have another reason to doubt that 'Hesperus' and 'Phosphorus' invoke the same persistence

conditions, as Gibbard claims. Indeed, if the persistence conditions that each name invokes depend upon the descriptive content that is associated with each name, there is no way to determine if the persistence conditions that 'Hesperus' and 'Phosphorus' respectively invoke are the same because, even if it is determinate that each name is associated with a particular descriptive content, it is indeterminate which specific descriptive content is associated with each name. In light of these considerations, it is clear that the No Explanatory Loss argument does not provide conclusive support for Gibbard's theory.

Reply to the Failed Rigidity Test Argument. After complaining that Kripke's test fails to support the thesis that proper names are rigid designators *simpliciter*, Gibbard (1976: 199) further states that, because the issue of whether names are rigid *simpliciter* or sortal-rigid cannot be solved by apealing to intuitions about whether names refer to mere entities or entities-*qua*-*something*, 'the matter has to be settled by working out rival systems and comparing their implications.' This is precisely what I intend to do here.

Are there any costs or disadvantages if we maintain, as Gibbard does, that names are sortal-rigid? I believe that there is a major problem associated with this position. Here it is. If names are sortal-rigid, this is presumably in virtue of the metaphysical fact that all objects that can be referred to using a name fall under some sortal concept. After all, if names are sortal-rigid, their rigidity has to be *de facto* because a designator is *de jure* rigid just in case the semantical rules of the language unmediately tie it to the object it refers via a stipulation and names denote according to Gibbard only through

the mediation of a sortal concept. But how plausible is it to hold that all objects that are referred to using a name must fall under some sortal concept? Although some philosophers (which I will call *sortalists*) have endorsed this position,[25] I think it is implausible in light of their own view about sortal concepts. Indeed, considering that sortalists typically characterize a sortal concept, following Lowe (2007: 524), as a concept that 'allows the singling out of an object in thought by a thinker only as an object of a specific sort', claiming that objects that are refered to using a certain name have to fall under some sortal concept amounts to claim that there are ways of singling out in thought these objects as objects belonging to specific sorts when we use names to denote them. But this is not true, since there are circumstances where a speaker can use a name n competently to refer to an object o without having a way of singling out in thought o as an object of some specific sort — i.e., without having what Evans (1982: 89) calls 'discriminating knowledge'.

In order to see this, let me first rehearse some plausible competence conditions for using names. According to Soames (2002: 65), some speaker S is a competent user of a name n for an object o iff S has a referential intention that determines o as the referent of n and S realizes that assertively uttering 'n is F' is to say of the referent of n that it 'is F'. Now, if these competence conditions are correct, asking whether a speaker can be a competent user of a name n without being able to single out in thought the referent of n as an

[25]Sortalism comes in many different versions. For instance, Dummett (1981: 179) associates it with a form of anti-realism in which the world is not divided *per se* into objects, but only in virtue of the sortal concepts that we use to categorize it. Wiggins (2001), on the contrary, holds that the world is already sliced up into objects, but that we cannot individuate them (in the epistemological sense of the term) unless they fall under some sortal concept. Given that these differences are irrelevant vis-à-vis the issue I want to address, I will ignore them here.

object of a specific sort boils down to ask two questions:

> (I) Can a speaker have a referential intention that determines o as the referent of n without having a way to single out in thought o as an object of a specific sort?
>
> (II) Can a speaker realize that assertively uttering 'n is F' is to say of the referent of n (i.e., o) that it 'is F' without having a way to single out in thought o as an object of a specific sort?

Let us consider (I). In order to tackle this question, it is important to spell out in some detail how speakers acquire a referential intention that determines a certain name n as the referent of an object o. Vis-à-vis this issue, Soames (*ibid.*) suggests two different ways: by picking up n from other speakers that used it as a name of o and intending to use it in the same way that they did or by being independently acquainted with o and introducing n as a name of o. Now, if we focus on the second way, which requires direct acquaintance with the relevant object, one may then rely on a traditional argument to answer (I) affirmatively. According to this traditional argument (Russell 1910: 119-120), having direct acquaintance of an object does not require having a way to single out the object in thought (or, using Russell's terminology, having an 'idea' of the object) because the opposite view 'leads at once to a vicious regress, since the relation of idea to object will have to be explained by supposing that the idea itself has an idea of the object, and so on *ad infinitum*.' In light of the abovementioned infinite regress argument and other parallel considerations,[26] I conclude that one

[26]Lowe (2007: 528) presents another compelling observation that supports the thesis

may provide an affirmative answer to (I) because a speaker can be directly acquainted with an object o and introduce a name n as a name of o without having a way to single out o in thought as an object of a specific sort.

Let us turn to (II). I believe that one may also provide an affirmative answer to this question in light of the following thought experiment. Suppose that an archeologist digging at an excavation site unearths an object that he is unable to identify.[27] The object in question has many distinct characteristics (e.g., it is red, disc-shaped, made of wood, etc.) that the archeologist perceives, but he is unable to determine on the basis of them what *sort* of object he found.[28] We may further suppose that the archeologist affectionately names his discovery 'Pakal', a term which he uses in conversations with his colleagues to describe to them how the newfound object is. Now, since the archeologist plausibly realizes when he assertively utters 'Pakal is red' that what he says is that the object he unearthed 'is red' (even if he does

that being directly acquainted with an object does not require an ability to single it out in thought as an object of certain sort which involves appealing to the cognitive capacities of animals: 'For it is evident that many non-human animals *perceive* objects in their immediate environment, even though it would be utterly extravagant to suppose that those animals are capable of *categorizing* those objects ontologically or grasping the relevant criteria of identity for those objects.' Fodor (1986) makes a similar point when he argues that simple systems such as paramecia that merely take in information do not have mental representations.

[27]This example is drawn from Campbell (2002: 70-71).

[28]An advocate of Gibbard's view might interject here the following objection: the unearthed object clearly falls under the concept *physical object*, which some authors (e.g., Xu 1997) take to be a genuine sortal concept. However, I am unpersuaded by Xu's arguments in favor of his thesis in light of the following consideration: although I agree that the concept *physical object* allows us to answer the *what is it?* question with respect to the item that the archeologist unearthed, it fails to provide determinate criteria of persistence for it. Considering this, the best alternative is to consider the concept *physical object* as a 'dummy sortal' —i.e., as a placeholder or a variable that stands for whatever item happens to be referred to (if any). For further discussion on dummy sortals, see Hirsch (1997) and Wiggins (1997).

not know how to single it out in thought as an object of a certain sort), it is clear that the second condition to use competently a name n to refer to an object o can also be met even we have no idea of what sortal concept o falls under.

To conclude, given that a necessary condition for names being sortal-rigid is that the objects that are referred to by these names fall under some sortal concept and that this is highly implausible because it entails there is a way to single them out in thought as objects of a certain sort when we use names to refer to them (and I have argued that this not necessarily the case), Gibbard's thesis that names are sortal-rigid involves a serious difficulty that Kripke's view is able to avoid.

Reply to the Cross-World Identities Worries Argument. The key element behind this argument is the stipulation that Gibbard makes —namely, the stipulation that there is some counterfactual situation W_1 in which Goliath and Lumpl are simultaneously brought into existence as they are in the actual scenario $W_@$ (namely, by sticking the two halves together) but in which, before the clay has the chance to dry, it is squeezed into a ball (an action which brings Goliath to an end but preserves Lumpl). If this stipulation is indeed legitimate, it should respect a plausible constraint that, according to Salmon (2005b: 222-223), all legitimate stipulations have to abide by: "Once certain stipulations have been made, one is automatically barred from making further stipulations that are not compossible with the initial stipulations under pain of 'stipulating' an impossible state of affairs."

Let us take a closer look at how Gibbard introduces the stipulation. Here is how he (1975: 191) initially characterizes W_1: 'For suppose that I had

brought Lumpl into existence as Goliath, just as I actually did, but before the clay had a chance to dry, I squeezed it into a ball. (...) Hence, Lumpl would not be Goliath, even though both existed.' The terms that Gibbard uses to describe W_1 (in particular, the use of the adjective *both*) strongly suggest that he treats Goliath and Lumpl as two distinct entities (i.e., Goliath≠Lumpl). This characterization stands in sharp contrast with the following passage, in which Gibbard (1975: 200) further characterizes W_1: 'After I made that thing, I held it in my hands and I could have squeezed it. If I suppose that I did squeeze it, I have stipulated as much about the identities of the things in that supposed situation as can be stipulated.' Indeed, since Gibbard's stipulation involves talking about *the thing* that he could have squeezed in W_1, he clearly considers Goliath and Lumpl as one and the same thing (i.e., Goliath=Lumpl).

Since W_1 involves conflicting stipulations, it cannot be used to argue, *pace* Gibbard, that cross-world identity worries still arise even after a counterfactual situation has been completely stipulated. I agree therefore with Gibbard that no definite answer can be provided to the question 'Which of the two (i.e., Goliath or Lumpl) is the one thing that, in the actual world, I made and then broke?' in W_1, but not because it does not make sense to ask about the counterfactual on-goings of a thing independently of how it is characterized. Rather, it is because the counterfactual situation W_1 is a clear case of overstipulation that results in an impossible state of affairs.

Reply to the Leibniz's Law Argument. The core of Gibbard's reply to this argument is that an advocate of CI can block the inference from L_1 to L_2, which requires the use of Leibniz's Law, by arguing that Leibniz's Law is

a law concerned with the genuine propreties that objects have (i.e., properties that the objects possess regardless of how they are characterized) and that the question of whether *de re* modal contexts such as '□(Lumpl exists → (_____=Lumpl)' expresses genuine properties or not is a point of contention between essentialists like Kripke and anti-essentialists like himself. In light of this, allowing the inference to go through requires us to assume that *de re* modal contexts express genuine properties, but this begs the question against anti-essentialism —and, consequently, against (CI).

Vis-à-vis this argument, I put forth the following reponse: allowing the inference from L_1 to L_2 to go through certainly requires assuming that modal contexts express genuine properties (and, consequently, essentialism about individuals), but Gibbard is not entitled to argue that this begs the question against his view because he himself needs to presuppose essentialism about individuals in order to get off the ground the Goliath-Lumpl scenario. In order to appreciate this clearly, let me first rehearse some general remarks that Salmon makes regarding how the problem of cross-world identity arises within the framework of anti-essentialism (which he labels 'anti-haecceitism'):

> On Anti-Haecceitism regarding individuals, possible worlds do not include specific individuals themselves. Instead they provide a structure and framework, given purely qualitatively, in which individuals are represented by means of individual concepts. It is not labeled which individual a given individual concept represents. For the Anti-Haecceitist, then, there is a special problem about how the individuals thus represented in distinct possible worlds are to be identified with, or distinguished from, one an-

other. (2005a: 208)

Though Salmon does not intend these remarks to describe specifically Gibbard's anti-essentialism, it is not difficult to see that they capture the metaphysical core of Gibbard's proposal. In fact, the Carnapian semantics that Gibbard presents to argue that the acceptance of CI and the concomittant theory of names entail no loss of expressive power clearly depend on the metaphysical picture that Salmon describes in the abovementioned passage. Now, if Gibbard is indeed committed to this metaphysical picture (and he is, under pain of depriving of his underpinnings the semantics he proposes), he is required to construe counterfactual situations as not including specific individuals. But that does not seem to be the case, as we are able to appreciate when we consider again the initial passage where he stipulates the existence of the counterfactual situation W_1:

> For suppose I had brought Lumpl into existence as Goliath, just as I actually did, but before the clay had a chance to dry, I squeezed it into a ball. (1975: 191)

As the passage shows, Gibbard supposes that it is he himself (and not some individual similar to him that is identified by fulfilling a certain qualitative condition) that squeezes the clay into a ball at W_1. Now, Gibbard may reply that my observation is not decisive since there are alternative ways to characterize W_1 that do not involve he himself squeezing the clay into a ball. Perhaps, as some ersatzists suggest,[29] what accounts at $W_@$ for the fact that

[29]For a thorough discussion of different types of ersatzism regarding possible worlds, see Lewis (1986), chap. 3.

Gibbard could have squeezed the clay into a ball is that at W_1 there is some entity e distinct from Gibbard that goes proxy for him (e.g., the name 'Allan Gibbard') and that satisfies the quantified statement (in which the quantifier is restricted to W_1):

> (QS) $(\exists_{W_1} x)(x$ brings Lumpl into existence as Goliath, just as Gibbard does in $W_@$, but before the clay has a chance to dry, x squeezes it into a ball).

However, even if we accept that it is not Gibbard himself that squeezes the ball at W_1, essentialism about individuals still crops up in the Goliath/Lumpl scenario in an inescapable way. In order to appreciate this, consider the following passage in which Gibbard further describes W_1:

> Take a possible world in which I squeeze Lumpl into a ball and suppose all the molecules involved are clearly identified. There are still two distinct things in that world, the statue Goliath which I destroy by squeezing, and the piece of clay Lumpl which survives the squeezing. (1975: 200)

According to the previous quote, Gibbard is not only committed to the existence of two distinct individuals at W_1 (namely, the statue and the lump of clay) but he also seems to be committed to something that destroys the statue and that merely reshapes the lump of clay —namely, the *squeezing*. Indeed, even though Gibbard does not explicitly state that he accepts an ontology involving events, his endorsement of the thesis that physical concrete things such as statues are built up from some fundamental entities

(among which he considers point-instants and particles as the most viable candidates) and his acceptance of a metaphysical view that involves temporal parts constitutes strong evidence that this is the case.[30] Now, assuming that Gibbard is committed to the existence of the squeezing of the clay, this event, which is presumably an individual,[31] has some its properties accidentally and some necessarily independently of how it is described. For instance, even if the squeezing of the clay into a ball could have been done using a person's feet rather than his hands, it necessarily involves the application of a force to the clay (regardless of whether we describe the event as a squashing, a compression, a constriction or a crushing). In addition, though the squeezing of the clay into a ball could have been done unintentionally (say, if the person holding the clay had squeezed it because of some involuntary hand spasms brought about by a neurological condition), it necessarily involves the piece of clay (and not, say, a hunk of metal).[32] Thus, since Gibbard is compelled to presuppose essentialism about a certain kind of individuals —namely, events— to get the Goliath-Lumpl scenario off the ground, it is patent that he has no solid grounds to block the inference from L_1 to L_2 in the Leibniz's Law argument against CI.

Reply to the No Diminished Expressive Power Argument. As we saw earlier, a key part of this argument is the Carnapian semantics that

[30]Quine (1960: 171), who also endorses a view of objects according to which they are made up from temporal parts, holds that 'physical objects, conceived thus four-dimensionally in space-time, are not to be distinguished from events (...) Each comprises simply the content, however heterogeneous, of some portion of space-time, however disconnected and gerrymandered.' Cf. also Goodman (1977), chap 4, sect. 1.

[31]For a classical discussion concerning the nature of events as individuals, see Davidson (1970).

[32]Consequently, I am in partial agreement with Kim's theory of events (1993: 48) according to which the constitutive substance of an event is essential to it.

Gibbard presents in order to accomodate modal discourse. Since individuals do not have modal properties *per se* but only relative to certain sortal concepts, Gibbard claims that, under pain of being unable to say many things we say in scientific talk and in daily life, one has to find an alternative way to interpret *de re* modal attributions to objects and talk that refers to dispositions —and this is precisely the function of the Carnapian semantics he introduces. Gibbard also maintains that, even though constants, variables and predicates have distinct references and ranges if they occur within a modal context (which is one of the consequences of the Carnapian semantics he adopts), this modification entails no loss of expressive power: we are still able to say the exactly the same things we said prior to the adoption of the Carnapian semantics.

The point in case that Gibbard discusses is the identity predicate '='. Although 's=c' expresses, if it occurs outside a modal context, an identity relation between s and c, it no longer expresses identity when it occurs inside a modal context. However, this is not problematic for Gibbard because, if they occur inside a modal context, 's' and 'c' no longer denote individuals but rather individual concepts. Thus, if 's=c' occurs outside a modal context, what it expresses is the identity of s and c and, if it occurs inside a modal context, it states that s and c are individual concepts of the same individual $\mathcal{I}$. According to Gibbard (1975: 206), this amounts to say that the individual of which s is an individual concept is identical to the individual of which c is an individual concept. If this is indeed true, the deployment of the Carnapian semantics to interpret modal discourse does not entail changing the subject: when we assert a claim of the form '$\phi(s{=}c)$' (where ϕ is a modal context), we

are still talking about the identity between the individual s and the individual c.

Unfortunately for Gibbard, saying that s and c are individual concepts of the individual $\mathcal{I}$ is not tantamount to say that the individual of which s is a concept is identical to the individual of which c is a concept. Indeed, in spite of the fact that it mentions the individual $\mathcal{I}$, the thesis that s and c are individual concepts of the individual $\mathcal{I}$ does not express a relation between an object and itself but rather a relation between concepts of different levels. In fact, what the thesis really states is that the two first-order concepts s and c fall under the second-order concept *is an individual concept of the individual* $\mathcal{I}$.[33] Now, if this is indeed the case, what motivated Gibbard to maintain that saying s and c are individual concepts of the same individual $\mathcal{I}$ boils down to say that the individual of which s is a concept is identical to the individual of which c is a concept?

[33]Let me present another problem, which illustrates the need to distinguish first-order relations between objects and second-order relations between first-order concepts. In a posthumous paper titled 'Comments on Sense and Meaning', Frege addresses, among several other problems, the following question: considering that $(x+1)^2=2(x+1)$ and $x^2=1$ have the same value for every argument they take, how can we express that the two concepts *less by 1 than a number whose square is equal to its double* and *square root of 1* (which respectively correspond to the two functions) have the same extension? After suggesting a particular notation that allows us to do this, Frege (1979: 121) warns us that what we have expressed is 'that second-order relation which corresponds, but should not be confused with equality (complete coincidence) between objects.' Frege is adamant elsewhere in his writings about the need to distinguish first-order relations between objects and second-order relations between concepts because not doing so leads to serious difficulties. In particular, Frege (1980: 46) suggests that, though Hilbert aims to provide precise definitions of points, lines and planes as determinate *objects* of a mathematical theory using the axioms he gives in *Die Grundlagen der Geometrie*, he fails to do so because, as he writes to Hilbert, "the characteristic marks you give in your axioms are apparently all higher than first-level; i.e., they do not answer the question 'What properties must an object have in order to be a point (a line, plane, etc.)?', but they contain, e.g., second-level relations, e.g., between the concept *point* and the concept *line*."

Here is a plausible answer: Frege, who was the first author to notice the importance of distinguishing sharply objects from first-order concepts (and first-order concepts from second-order concepts), noticed that, although the successful deployment of the mathematical categories of argument and function to provide an analysis of natural language statements such as 'Caesar conquered Gaul' calls for an ontology that admits both objects and concepts,[34] concepts are rather peculiar because, strictly speaking, we cannot really refer to them. Since concepts are akin to functions for Frege (as they are also for Gibbard, who identifies an individual concept with a function that assigns to each possible world in a set an individual in that world), they are *predicative*. Consequently, in light of their predicative nature, concepts cannot be denoted unless, using Frege's (1892a: 197) words, one introduces certain objects that go proxy for them (e.g., the object denoted by the name 'the concept horse' goes proxy for '(...) is a horse'). However, undertaking this maneuver requires extreme carefulness since it is very easy subsequently to conflate talk about concepts with talk about objects —which is what happens to Gibbard when he asserts that saying the concept s and the concept c are concepts of the same individual is tantamount to say that the individual of which s is a concept is identical to the individual of which c is a concept. Thus, in light of the fact that the No Diminished Expressive Power Argument trades on the abovementioned conflation, I conclude then that Gibbard is not entitled to claim that the Carnapian semantics he adopts enables him to preserve all the things we usually say both in scientific talk and in daily life.

[34]For a detailed discussion of how Frege's semantic views have important implications for his ontology, see Wells (1951).

3.2.3.2 A new argument against CI

In the previous subsection, I offered systematic replies to the main arguments that Gibbard presents to support (CI). In the current subsection, I develop a new argument against CI which shows that a key presupposition that underlies Gibbard's view (namely, his anti-essentialism about individuals) is incompatible with his endorsement of the thesis (1975: 194) that cross-world identities between individuals are genuine identities, albeit relativized to sortal concepts .

I have previously argued that, in light of the anti-essentialism about individuals that Gibbard espouses, what plausibly accounts according to him for the fact at $W_@$ he could have squeezed the clay into a ball is not that he himself is present at W_1 and squeezes the clay into a ball but rather than some entity that goes proxy for him satisfies QS. Likewise, what presumably accounts according to Gibbard for the fact that at $W_@$ Goliath would have been destroyed whereas Lumpl would have persisted had the clay been squeezed is that certain entities that go proxy for Goliath and Lumpl satisfy certain conditions at W_1 —namely, they are respectively selected as values at W_1 by the individual concepts '(...) is a clay statue that is destroyed by the squeezing' and '(...) is a lump of clay that is reshaped by the squeezing'. Let us call these representatives or surrogates of the actual individuals at W_1 *counterparts*. At this point, the following question naturally arises: what exactly are these counterparts?

There seems to be only two available answers to the previous question. The first one, which Lewis (1986) endorses, identifies counterparts with concrete individuals that are constituents of possible worlds. Now, considering

that Lewis takes possible worlds to be entities that are of kind with our world —which he describes as a very inclusive thing that encompasses not only all the entities in our immediate surroundings but also any other entity with which we are spatiotemporally related, no matter how remote in space or time it is— but that are spatiotemporally disjoint with respect to it, his version of counterpart theory is tied to his modal realism. But, since modal realism is often decried as an implausible view (if not outright paradoxical), other philosophers have developed versions of counterpart theory that are independent of modal realism. For instance, Stalnaker (1986) endorses a version of counterpart theory embedded in an actualist framework in which counterparts are identified with actual but unexemplified properties of objects.[35]

Thus, in virtue of his anti-essentialism about objects, Gibbard is then compelled to accept that the task of representing *de re* the modal properties of actual objects is undertaken by counterparts, and that these are either Lewisian *possibilia* or some kind of actual abstract entities (e.g., properties or linguistic constructs). However, neither alternative allows him to maintain that there are genuine cross-worlds identities between individuals. In order to appreciate this, let us consider first the view that the counterparts are Lewisian *possibilia*. If this is the case, then it is patent that, even if we concede that the fact that o is possibly F is identical with the fact that o has some counterpart o' that is F,[36] o and o' are distinct individuals to the extent that they are part of distinct worlds and that individuals are world-

[35]In addition to Stalnaker, other authors who also embrace an actualist version of counterpart theory in which counterparts are identified with actual abstract entities that represent possible objects include Heller (1998) and Sider (ms).

[36]See Lewis (1986: 196).

bound according to Lewis. In other words, even though we may grant that Lewis' view allows us to provide a genuine analysis of what it is for an individual to have a modal property, the price to pay is that we cannot have genuine cross-world identity between individuals. Thus, construing counterparts as Lewisian *possibilia* precludes Gibbard from maintaining that there are genuine cross-world identities between individuals.

Let us turn now to the second alternative. According to it, counterparts are abstract entities of some kind that exist in the actual world and that represent *de re* the modal properties and the counterfactual on-goings of actual objects. But can an actual abstract entity e represent *de re* a possible object o' in such a way that it allows a genuine cross-world identity between o' and some actual object o? For Stalnaker, the answer is yes: we are just required to accept certain plausible theses to see how this can be achieved. These theses, which are strongly motivated within an actualist framework as he shows, are the following:

> First, a distinction between possible individuals themselves and their essences or representatives in the domains of possible worlds and, second, the rejection of a world-independent concept of identity. (...) Since, according to actualism, there are no merely possible individuals, we need something other than possible individuals to represent the possibility that there be things other than those that actually exist. And since, according to actualism, the actual world is not one place among others but is the only place there really is, real truth is truth at the actual world. (1986: 133)

Once we accept a distinction between possible individuals and their surrogates and the thesis that all statements (including cross-world identity statements) are, if true, true at the actual world $W_@$, we have, according to Stalnaker, the tools needed to answer affirmatively the abovementioned question. To illustrate this, suppose that there is an abstract entity e_1 (e.g., an unexemplified actual property) that represents *de re* Goliath's being destroyed by the squeezing of the clay at W_1. If we further suppose that e_1 represents Goliath-at-W_1 at the actual world $W_@$ via a certain counterpart relation C that holds between e_1 and some particular property e that is uniquely exemplified by Goliath-at-$W_@$, Stalnaker claims that one may provide an account of how this counterpart relation C underpins a genuine transworld identity in the following terms: Goliath-at-W_1 is identical to Goliath-at-$W_@$ (at $W_@$) iff e_1 bears the relevant counterpart relation C to e (at $W_@$).

Unfortunately, a thorough look at Stalnaker's proposal reveals that it cannot deliver cross-world identities that are genuine identities —and this makes his proposal unfit to be used by Gibbard. In order to see this, let me rehearse briefly how Stalnaker answers a potential objection to his proposal that stems form some observations made by Salmon. While commenting Kaplan's (1975: 722) characterization of haecceitism as the view according to which 'it does make sense to ask —without reference to common attributes or behavior— whether this is the same individual in another possible world', Salmon (2005b: 127fn18) observes that 'at least one version of haecceitism may be described as the view that the logic of possible world discourse includes a binary *absolute* concept of genuine identity in addition to its ternary *world-relative* concept of identity.' Since Stalnaker accepts the actualist the-

sis that all statements are, if true, true relative to the actual world, he is concerned that someone might argue, following Salmon, that intraworld identity statements do not really express genuine identity (which, as Salmon suggests, should be an absolute, world-independent relation). To block this potential objection, Stalnaker responds in the following way:

> Suppose we reject the assumption that there is an identity relation that is absolute in Salmon's sense. Can we still explain the intraworld or world-relative identity relation as the relation of 'being the very same thing', irrespective of ways of conceptualizing the things identified? There is no reason why we cannot: [intra-world] identity is the binary relation whose extension is, in any possible world w, the set of pairs $<d, d>$ such that d is in the domain of w. This is surely identity. (1986: 132)

In light of the fact that Stalnaker characterizes in the abovementioned passage the intra-world identity relation in terms of a set of pairs of the form $<d, d>$, it is clear that intra-world identity is for him nothing but self-identity. Now, since self-identity is primitive in the sense that nothing is required to explain why, for any d, d is identical with d,[37] it seems that Stalnaker endorses the view that intra-world identity (which is 'surely identity', as he puts it) is primitive. But this clearly conflicts with the account of cross-world identity that he endorses. Indeed, considering that a cross-world identity between individuals holds for Stalnaker in virtue of the fact that a counterpart relation obtains between the representatives of these individuals,

[37]For a persuasive argument for this claim, see Jubien (1996: 350).

it is patent that cross-world identity is not primitive in his view since it is explained in terms of another relation. Putting this in other terms, if Gibbard appealed to Stalnaker's proposal to explain how the thesis that cross-world identities are genuine identities is compatible with CI, he would face an intractable difficulty: explain how both intra-world identities and cross-world identities can be instances of one and the same relation in light of the fact that the former are primitive and the latter not.[38] Consequently, given that there is no way to reconcile one of the central claims underpinning CI (i.e., the anti-essentialism about individuals) with Gibbard's endorsement of the thesis that cross-world identities are genuine identities, I conclude then that Gibbard's views are internally inconsistent and constitute no real challenge to Kripke's defense of the necessity of identity.

3.3 Wreen's challenge to the necessity of true identity statements involving names

I considered in the previous section Gibbard's arguments for the thesis that, *pace* Kripke, there are certain identity statements involving exclusively names that are true but only contingently so, and I showed that all of them are either flawed or inconclusive. In this section I turn to a different objec-

[38]There are two main strategies to handle this difficulty. The first one, which Adams (1979) endorses, consists in holding that cross-world identities are primitive. The second one consists in claiming that intraworld identities are not primitive and must be explained in terms of some other relation. However, none of the strategies are available to Gibbard if he wants to maintain his anti-essentialism about individuals while using Stalnaker's proposal to account for cross-world identity. Indeed, his anti-essentialism about individuals is incompatible with the primitiveness of cross-worlds identities and Stalnaker's proposal is at odds with the thesis that intraworld identities are non-primitive.

tion that targets (iii), which has been raised by Wreen (1998), and I show that this objection can be safely dismissed by an advocate of Kripke's views because the argument that supports the objection trades on a series of subtle confusions and contradictory assumptions. Here is a brief summary of how I accomplish this. First, I present in §3.3.1 Wreen's argument against (iii). Having done that, I take a closer look in §3.3.2 at several inferences upon which the argument depends and I show that, while some trade on assumptions that are at odds with each other, others require accepting distinct (and conflicting) interpretations of certain principles that Wreen relies on. Finally, I argue in §3.3.3 that the main reason for which Wreen's argument founders is that it fails to respect a very important distinction introduced by Kaplan in his discussion of rigid designators —a distinction which, if not respected, creates serious problems.

3.3.1 Wreen's argument

The first two premises of Wreen's argument state some uncontroversial facts about what the names[39] 'Hesperus' and 'Phosphorus' respectively refer to. Since 'Hesperus' refers to Hesperus (which is Venus) and 'Phosphorus' refers to Phosphorus (which is also Venus), we have:

(1) 'Hesperus' designates Hesperus (i.e., Venus)

and

[39]In order to sidestep potential problems stemming from the fact that the same name may be used to refer to different objects in different circumstances (e.g., 'Aristotle' may be used to refer to Aristotle of Stagyra or Aristotle Onassis), Wreen specifies that he considers names to be *token-names* (i.e., individual occurrences of the same *name-type*) that denote a unique object. This move is considered sympathetically by Kripke (1980: 8)

(2) 'Phosphorus' designates Phosphorus (i.e., Venus).

From (1) and (2), the following thesis can be unproblematically derived:

(3) 'Hesperus' and 'Phosphorus' designate the same object.

If, in addition to (3), we also accept (as it seems plausible) what Wreen calls the *Disquotation Principle* (DP) which states that if two singular terms 'A' and 'B' denote the same object then A is B, we may obtain from (3) and DP:

(4) Hesperus is Phosphorus.

After establishing (4), Wreen reminds us that Kripke (1980: 91-97) presents a picture that aims to provide an account of how names refer to objects. According to this picture, a speaker can use a name n to refer to an object o only if the speaker's utterance of n is causally grounded by o (i.e., if o lies at one end of a certain causal chain which has at its other end n).[40] For Wreen, the core of Kripke's picture is captured by the following thesis:

(KTR) A token-name n designates (or refers to) an object o if and only if it is o that causally grounds that token.

[40]The picture that Kripke presents is much more complicated than what this brief statement expresses. In particular, Kripke mentions that the causal grounding of a name N by an object o can take place in a variety of ways. For instance, if the speaker is initially introducing a name n to refer to o in a baptism-like ceremony, the speaker may use a definite description or an act of ostension to fix o as the reference of n. If the speaker has, on the contrary, learned the name n from another individual who used it to refer to o, he must intend to use it with the same reference than the individual from whom he picked it. Since these details are irrelevant to Wreen's argument, I will ignore them here.

Using KTR and the plausible thesis that *causal connections are contingent* (which he dubs CC), Wreen argues that the two following claims can be respectively obtained from (1) and (2):

(5) (1) is contingent

and

(6) (2) is contingent.

If (1) and (2) are contingent, the causal chain that actually ties 'Hesperus' to its referent and the causal chain that actually ties 'Phosphorus' to its referent could fail to exist, as it would be the case if the names were assigned to distinct objects. Thus, since it is a contingent fact that the two chains are grounded by the same object, we have:

(7) (3) is contingent.

Now, Wreen further argues that if we assume that not all the counterfactual situations are under consideration but merely those in which Hesperus and Phosphorus exist —a restriction that Kripke (1980: 110) seemingly endorses since he admits that "if we wish to be somewhat more careful we could replace the statement 'Hesperus is Phosphorus' by the statement 'If Hesperus exists then Hesperus is Phosphorus', cautiously taking only the latter to be necessary"— and also further restrict our attention to those in which our words 'Hesperus' and 'Phosphorus' exist,[41] we then have

[41]Even though Wreen (1998: 322) characterizes this restriction as 'admittedly artificial but relatively innocuous', this restriction has some rather profound consequences which seriously undermine his argument. For further discussion of this, see §3.3.2 below.

(8) (3) is logically equivalent to (4)

because, in virtue of the abovementioned restrictions, (4) entails (3) and, since we have accepted previously DP, (3) entails (4). Finally, if we also accept, in addition to (3) and (8), the following principle (which seems unquestionable):

> (MS) Logically equivalent statements have, and must have, the same modal status.

we can derive the following conclusion, which is at odds with (iii):

(9) (4) is contingent.

3.3.2 Some problems with Wreen's argument: inconsistent assumptions and conflicting interpretations

Since the conclusion of Wreen's argument states that the identity statement 'Hesperus is Phosphorus' is contingent, it constitutes, if the argument is sound, a counterexample to Kripke's thesis that true identity statements between names are necessary. If an advocate of Kripke's is to reject the conclusion, he must find some fault in Wreen's argument. In the present subsection, I take a closer look at the argument and I show that it exhibits some problems that cast doubts on its soundness.

Let us consider steps (5)-(9), which Wreen views as the most controversial part in his argument.[42] With regard to these steps, I want to bring attention

[42]Concerning steps (1)-(4), I grant Wreen's observation that they are unobjectionable and, accordingly, they are not discussed here.

to a couple of problems that come to light when we focus on what is required to maintain that (1) is contingent in light of the restrictions that Wreen introduces to establish (8) on the one hand and when we consider carefully the inference from (8) and MS to (9) on the other hand.

Intuitively, (1) is contingent iff it is possible that 'Hesperus' does not refer to Hesperus. Now, there are several distinct types of possible scenarios at which 'Hesperus' does not refer to Hesperus. For instance, considering that Hesperus is not a necessary existent, there are possible scenarios at which Hesperus does not exist and 'Hesperus' is causally grounded by another object or not at all. There are also other possible scenarios at which Hesperus exists but 'Hesperus' does not (i.e., scenarios at which people speak a non-standard version of English where 'Hesperus' is absent). Finally, there are possible scenarios at which both Hesperus and 'Hesperus' exist, but 'Hesperus' is grounded by a different object (say, Mars). Since Wreen maintains that the possible scenarios under consideration in his argument are only those of the third kind, it is clear that he is committed, in addition to genuine cross-world identities between objects, to genuine cross-world identities between names. But how can one endorse this thesis (as Wreen does) if one does not have a concommitant view about what a name consists in? In light of the need to have a view about what names consist in, Wreen presents three distinct options which, as far as he can tell, are the only available ones:

(N_1) A name is a symbol associated to a descriptive sense.

(N_2) A name is just an uninterpreted symbol.

(N_3) A name is a symbol associated to a referent.

Even though Wreen is cautious enough not to endorse explicitly any of these views (or to attribute any of them definitively to Kripke), he makes certain remarks in his paper that reveal an underlying tension that is fatal to his argument. Indeed, after he introduces N_2 (which he calls the Bare Inscription Theory), Wreen strongly suggests that it is only if this view of names is assumed that we can derive (5) from (1) using KTR and CC:

> On the Bare Inscription Theory, that a given token-name, 'Hesperus', designates Hesperus —Hesperus the planet— certainly seems contingent. For *no matter what* theory of reference for names is coupled with the Bare Inscription Theory, that theory of names is conceptually independent of the identity of the names (token-names) that it functions to map onto objects (referents). (1998: 326)

If this view of names is required to derive (5), it should be held uniformly throughout Wreen's argument for the sake of its consistency. But this is not the case, as a closer look at the restrictions that Wreen introduces to justify (8) shows. Recall that Wreen maintains that he does not consider in his argument all the counterfactual situations, but that he restricts his attention to those in which Hesperus the planet (and Phosphorus the planet) exists and our names 'Hesperus' and 'Phosphorus' also exist. It is patent that Wreen needs to introduce this restriction to be able to obtain (3) from (4) and, consequently, to derive (8), but this restriction is far from innocuous. Indeed, when Wreen restricts his attention to a certain class of counterfactual situations (namely, those in which our object Hesperus —which is the same as our object Phosphorus— exists and our names 'Hesperus' and 'Phosphorus'

also exist), he is implicitly accepting a view of names different from the Bare Inscription Theory —a view which seems to be something like N_3 (which he calls the Referent Theory). Consequently, given that Wreen requires to presuppose a certain theory about what names are to make the inference from (1) to (6) go through and that he needs to assume a different theory to justify (8), his argument against (iii) founders in virtue of the fact that it presupposes at different stages conflicting theories about the nature of names.[43]

Let us now turn our attention to the inference from (8) to (9). Wreen justifies the inference by appealing to a certain principle that governs the modal status of logically equivalent statements. Concerning this principle, I argue further down that, even though Wreen takes MS to be unobjectionable because he presumably construes it in a way that makes it trivially true, interpreting MS in this way makes it unsuitable to be used to derive (9). In order to appreciate this, it is important to observe first that the very formulation of MS involves several terms (which I highlight using italics) that invoke modal notions: 'logically equivalent statements not only have, but *must* have, the same *modal* status.'

Once we notice this, the following question arises: which kind of modality do these terms refer to? Since Wreen takes MS to be an unobjectionable principle (in fact, he does not even bother to present any arguments for it), it is plausibly a kind of modality that makes MS trivially true. But what would that modality be? Here is a suggestion: since the principle aims to govern logically equivalent statements, perhaps the modality in question is

[43]Fitch (2004: 108-109) articulates a reply to Wreen's argument similar to the one I present here.

logical modality (i.e., modality that arises from the laws of logic and the meaning of logical constants). Indeed, if we interpret MS in terms of logical modality, it is then clear that the principle is unobjectionable. To illustrate this, consider two formulas of propositional logic that are logically equivalent such as P∨¬P and ¬(P∧¬P). Under the abovementioned interpretation of MS, it is unobjectionable that it is logically necessary that both formulas have the same modal status (i.e., both are logically necessary) precisely because the meanings of the connectives '¬', '∨' and '∧' are such that, on one side, P∨¬P comes out as true under any truth-value assignment to P and, on the other side, both formulas have the same truth-value for any truth-value assignment to P.[44]

Let us grant, on the basis of the considerations made in the previous paragraph, that Wreen is required to assume that the notion of modality underlying MS is that of logical modality to make MS unobjectionable. However, if this is the case, he cannot use MS in conjunction with (8) to obtain (9) because the notion of logical equivalence between statements that is employed in (8) is not the one that underlies MS. To see this clearly, it is crucial to bear in mind that Wreen is compelled to restrict the class of counterfactual situations he considers in his argument to justify (8): he only takes into account those counterfactual situations in which both the object Hesperus (and Phosphorus) and our names 'Hesperus' and 'Phosphorus' exist. When he enacts this restriction, Wreen builds into the notion of logical equivalence used in (8) a certain modal component, but that modal component is not of the same kind as the one that underlies the notion of logical equivalence used

[44]For a historical interpretation of the notion of logical equivalence between statements in terms of logical modality, see Carnap (1947: 11).

in MS. Whereas the modality that underlies MS is, as I argued previously, *logical* (that is, if we want to hold on to the idea that MS is a trivial, unobjectionable principle), the modality underlying the notion of logical equivalence used in (8) is *metaphysical* because it stems from a metaphysical contraint that Wreen imposes on the counterfactual situations under consideration. In virtue of this, it seems that we cannot use MS to obtain (9) from (8) since the notion of logical equivalence that is implicitly assumed in MS is not the notion that is deployed in (8). This casts a serious doubt on the soundness of Wreen's argument.

3.3.3 Wreen's argument and Kaplan's distinction between contexts and circumstances of evaluation

Even though the problems considered in the previous subsection raise major qualms with respect to the argument, a sympathizer of Wreen might reply that they are not decisive. Perhaps, he might argue, there is some view about names, different from the three that Wreen considers, that allows us to justify the inference from (1) to (5) while upholding (8), and perhaps there is a notion of logical equivalence that guarantees the unobjectionable character of MS and is consistent with the restrictions introduced to defend (8).[45] I confess that I am not persuaded at all by these considerations. However, even if I conceded the replies made by Wreen's sympathizer, there is an additional

[45] Wreen (1998: 323) makes such a suggestion when he writes that 'a new concept, *logical equivalence*$_1$, can be coined for the nonce and used in (8) and MS to yield (9)'. To this I respond that, unless there is some reason to accept this new concept of logical equivalence that is different and independent from the need to defend Wreen's argument, this new concept turns out to be a mere *ad hoc* stipulation.

objection that can be raised —an objection which, as I show further down, undermines decisively Wreen's argument.

Wreen argues that (1) is contingent on the basis that the fact that 'Hesperus' designates Hesperus rests on the existence of a certain causal chain and that the existence of this causal chain is contingent. Now, Kripke (1980: 72) might certainly grant that (1) is contingent given that he writes, commenting on another parallel example, that "it is not trifling at all to be told that Socrates was called 'Socrates'. If this is any kind of fact, it might be false." But acknowledging that a certain name (say, 'Socrates') could be used to refer to an object different from the one it is actually used to refer to (namely, Socrates) does not entail that the reference of the name varies across counterfactual situations. Kripke emphasizes clearly this when he considers what occurs when we stipulate a scenario at which people use differently our names (e.g., assigning to them referents distinct from the ones we actually assign) and argues that the contingency of the fact that a certain name n is actually used to denote an object o does not conflict with or undermine the thesis that proper names are rigid designators:

> We say (...) 'suppose we had been using English in a non-standard way.' Then we are describing a possible world or counterfactual situation in which people, including ourselves, did speak in a certain way different from the way we speak. But still, in describing that world, we use *English* with *our* meanings and *our* references. It is in this sense that I speak of a rigid designator having the same reference in all possible worlds. (1980: 77-78)

As this passage makes clear, Kripke recognizes that there is a sense in

which it is contingent that 'Hesperus' designates Hesperus —and this sense corresponds to a scenario in which we use non-standardly the English word 'Hesperus' to designate, say, Mars. However, this is not at odds with the thesis that proper names are rigid designators because, according to Kripke, the thesis concerns proper names *as we use them in the actual world*. Thus, Kripke would agree with Wreen's assertion that (1) is contingent on the grounds that the world might be different from what it actually is (in particular, it might be such that we use 'Hesperus' to designate Mars at it), but he would also observe that this does not contradict the claim that 'Hesperus' is a rigid designator because what he means by this claim is that, granting that the actual world is such that we use 'Hesperus' to designate Hesperus at it, 'Hesperus' designates Hesperus with respect to all counterfactual situations.

Kripke's remarks in the aforementioned paragraph hint at two distinct ways of conceiving possible worlds when we use them as tools to evaluate the semantic value of expressions. Though Kripke himself does not provide a systematic treatment of the suggestion underlying his words, other authors have developed more elaborate proposals aimed at capturing what Kripke had in mind.[46] In particular, one such proposal is presented by Kaplan, who fleshes out the distinction in the following terms:

> (...) we must distinguish possible occasions of *use* —which I call

[46]For instance, Stalnaker's (1978) two-dimensional matrices, which are used to represent *propositional concepts* (i.e., functions from ordered pairs of possible worlds to truth-values), were developed with the purpose of showing how, in addition to the fact that the truth-value of the proposition expressed by the utterance of a sentence depends on the facts (here, the expression 'the facts' is taken to include the way we actually use language as fixed), there is a second way in which the facts determine the truth-value of the proposition —a way in which, to use Stalnaker's words, it is assumed that 'it is a matter of fact that an expression has the content that it has.'

> *contexts*— from possible circumstances of *evalaution* of what was said on a certain occasion of use. Possible circumstances of evaluation I call circumstances or, sometimes just, *counterfactual situations*. A directly referential term *may* designate different objects when used in different *contexts*. But when evaluating what was said in a given context, only a single object will be relevant to the evaluation in all circumstances. (1989a: 494)

Prior to discuss how Kaplan's distinction can be used to show that Wreen's argument founders, let me introduce some clarificatory observations regarding the place that the distinction occupies in Kaplan's theory of demonstratives and the potential applicability of the distinction to the case of proper names. To grasp accurately the role of the distinction in Kaplan's theory, it is important to remember that the main objective of Kaplan's arguments in *Demonstratives* is to show that what he calls the Direct Reference Theory of Demonstratives provides the best model for demonstratives.[47]

Here is very roughly how he proceeds to show this. First, he (1989a: 490) points out that two notable semantic features of any demonstrative is that 'the referent is dependent on the context of use and that the meaning of the word provides a rule that determines the referent in terms of certain aspects of the context.' Having done that, he (1989a: 494) presents us with a picture that illustrates how directly referential expressions function. According to this picture, propositions are considered to be structured entities

[47]Kaplan claims that, in addition to demonstratives, the theory that he presents is also the best available model for indexicals. For the sake of simplicity, I try to focus solely on the case of demonstratives here but I take his main conclusions to apply equally well to demonstratives and indexicals alike.

similar to the sentences that express them and, if a sentence contains the occurrence of a singular term, it is assumed that there is a corresponding constituent in the proposition expressed. On the basis of this picture, Kaplan claims that, whereas the propositional constituent that corresponds to a non-directly referential singular term is a complex (which often involves a reference-determining mechanism and whatever happens to be determined by it in different contexts), the propositional constituent that corresponds to a directly referential singular term is just its referent. With the notion of a directly referential expression in place,[48] Kaplan then introduces the following two theories for demonstratives:

> Fregean Theory of Demonstratives (FTD): the reference of a demonstrative, which is the object picked out by the demonstration associated to the demonstrative, is determined by the *sense* of the demonstration, which is a reference-fixing rule. Thus, granting that a sentence involving an demonstrative expresses a proposition, the propositional component corresponding to a demonstrative is a certain reference-fixing rule that determines different individuals in different contexts.

[48] In addition to this characterization of a directly referential expression (which is given in terms of its contribution to propositions expressed by sentences that contain it), Kaplan also introduces a second characterization according to which a directly referential expression is one which picks out its referent directly without the mediation of an individual concept. Martí (2003) has recently argued that both characterizations of direct reference (which she calls *propositional* and *Millian* respectively) are not equivalent and that it is only if we consider the propositional notion of direct reference that we can defend many theses usually held about directly referential expressions (e.g., the thesis that directly referential expressions are rigid designators). Following Martí, I accept the distinction and consider here only the propositional characterization of direct reference to avoid complications.

> Direct Reference Theory of Demonstratives (DRTD): the reference of a demonstrative, which is the object picked out by the demonstration associated to the demonstrative, is given independently of any reference-fixing rule. Thus, granting that a sentence involving an demonstrative expresses a proposition, the propositional component corresponding to a demonstrative is just the individual denoted by the demonstrative, which is loaded into the proposition without the intervention of any reference-fixing rule.

After introducing the two theories, Kaplan presents us with an argument that shows that DRTD provides a better account of demonstratives than FTD. The gist of the argument consists in considering, given a certain context, what is the truth-value of a certain proposition expressed by a sentence involving a demonstrative under certain counterfactual circumstances. The sentence that Kaplan gives as an example is the following:

> (S) He now lives in Princeton, New Jersey.

Kaplan maintains that if S is uttered while pointing at Paul (who lives in Princeton, New Jersey), the proposition that S expresses (which he calls 'Pat') is true. After mentioning this, he (1989a: 512) considers what would happen if he were to utter S while pointing to Charles at a certain counterfactual scenario in which Paul and Charles (who both live outside New Jersey) had disguised themselves as the other respectively and had switched places. Kaplan argues that, if this had happened, the proposition that would have been expressed by the utterance of S would have been false but it would

not have been Pat —it would have been a different proposition, which he calls 'Mike', since it would have involved Charles rather than Paul. Kaplan further argues that Pat is true in all counterfactual circumstances of a context in which Paul resides in Princeton. Using this distinction between Pat and Mike, Kaplan then claims that DRTD is a better theory of demonstratives than FTD because, when we consider what are the truth-conditions for the proposition that S expresses at some counterfactual situation, FTD gets things wrong.

Indeed, according to DRTD, the proposition expressed by S at the counterfactual situation (namely, Pat) is false because the individual relevant for the evaluation is the individual that is actually demonstrated (namely, Paul) whereas, according to FTD, the proposition expressed by S at the counterfactual situation (Mike) is false because the individual relevant for the evaluation is the individual that would have been demonstrated had the counterfactual situation taken place (namely, Charles). Intuitively, this gets things wrong because, if our goal is to determine what are the truth-conditions of the proposition that is actually expressed by S at some counterfactual scenario, it is deeply misleading to consider what *would have been* the proposition expressed by S if that counterfactual scenario had obtained.

As my brief presentation of Kaplan's argument shows, the distinction between contexts and circumstances of evaluation is crucial because the distinction between Pat and Mike, which is a key element in the argument to show that DRTD is a better theory than FTD, depends upon it. Moreover, the distinction is of paramount importance because without it we are unable to account consistently for certain intuitions and, more importantly,

not respecting it lands us in serious problems. To appreciate the first point, consider the following example from Kaplan (1989a: 514), which consists in someone uttering (very slowly, of course) the following sentence while pointing to Venus at two distinct times of the day:

> (T) That [pointing to Venus in the morning sky] is identical with that [pointing to Venus in the evening sky].

When we are asked to determine the modal status of T, we are pulled by conflicting intuitions. On one side, we feel compelled to maintain that T is contingent because the act of demonstration performed in the evening could have picked out Mars rather than Venus. On the other side, we also feel compelled to maintain that T is necessary because, granting that the two acts of demonstration associated to the demonstratives pick out Venus, there is no counterfactual situation in which Venus is not Venus. Kaplan's distinction between contexts and circumstances of evaluation allows us to reconcile these two conflicting intuitions in the following terms: it enables us to hold that T expresses in other contexts a proposition different from the one it expresses in the actual context (a proposition that is false at the actual context), and it also enables us to hold that, granting that T expresses a certain proposition in the actual context, the proposition that T expresses is necessary because it is is true in all circumstances of evaluation of the actual context.

Concerning the second point, Kaplan (1989a: 509) vividly illustrates the difficulties that arise if we do not respect the distinction by asking us to consider the following sentence:

(N) $\Box$ I am here now.

Kaplan remarks that, if we do not accept the distinction between contexts and circumstances of evaluation, any utterance of N by him or anybody else turns out to express a necessary truth —and this is clearly false because, when N is uttered by Kaplan in Los Angeles on December 4, 2009, what N expresses is the proposition that Kaplan is in Los Angeles on December 4, 2009 and this proposition is clearly contingent. Consequently, given that any failure to respect the distinction creates serious difficulties, we have strong reasons to accept it.

Having highlighted the importance of Kaplan's distinction between contexts and circumstances of evaluation to argue that demonstratives are directly referential devices, I turn now to assess the potential applicability of the distinction to the case of proper names. Now, in the last section of *Demonstratives*, Kaplan expresses some reservations with respect to the project of using the technical apparatus he presents (which includes the distinction between contexts and circumstances of evaluation) to the analysis of sentences containing proper names. Although they are quite similar to demonstratives and indexicals in certain respects, Kaplan suggests that there are certain features that prevent extending his apparatus to proper names:

> The contextual feature which consists of the causal history of a proper name expression in the agent's idiolect seems more naturally to be regarded as determining what word was used than fixing the content of a single context-sensitive word. Although it is true that two utterances of 'Aristotle' in different contexts may

> have different contents, I am inclined to attribute this difference to the fact that distinct homonymous words were uttered rather than a context-sensitivity in the characer of a single word 'Aristotle'. (...) In this, proper name words are unique. They have the direct reference of indexicals, but they are not context-sensitive. (1989a: 562)

As the previous passage states, the main qualm that Kaplan has with respect to the eventual extension of his apparatus to proper names is that, even though they are directly referential, proper names are not context-sensitive and, consequently, the apparatus (which was developed precisely to deal with demonstratives and indexicals, which are context-sensitive expressions) is not well-suited for them. *Pace* Kaplan, I argue further down that his apparatus (and, in particular, the distinction between contexts and circumstances of evaluation) can be extended to proper names.

Before presenting the argument that I rely on to establish this, let me briefly point out a strategy that one might use to try to extend Kaplan's apparatus to proper names —a strategy that I do not endorse because I think it is mistaken. Considering that the problem that Kaplan stresses is that his apparatus cannot be extended on proper names because they are not context-sensitive, one may attempt to show that Kaplan's remarks about proper names are erroneous and that they really are indexicals (i.e., context-sensitive expressions of a certain type). Even though some philosophers have explored this path,[49] I think that there is a serious problem that this proposal

[49]For some elaborate defenses of the indexical character of proper names, see Almog (1981) and Pelczar and Rainsbury (1998).

faces which stems from the fact that names do not have a certain feature that characterizes indexicals. Indeed, even though many expressions (including indexicals) display context-sensitivity, the type of context-sensitivity that indexicals exhibit is distinctive, as Perry notes in the following passage:

> Sometimes we use context to figure with which meaning a word is being used, or which of several words that look and sound alike is being used, or even which language is being spoken. These are *presemantic* uses of context. In the case of indexicals, however, the context is used *semantically*. It remains relevant after the words, meanings and language are known; the meaning directs us to certain aspects of context to determine what is designated. (2000: 314)

To illustrate the distinction that Perry has in mind, consider the sentences 'The woods are beautiful' and 'I like pretzels'. Both are context-sensitive but, whereas in the case of the first sentence the context is used to determine which meaning of the expression 'the woods' is being used (and this, in turn, determines whether what is described is a certain forest or some pieces of timber in a carpentry shop), in the case of the second sentence the meaning of 'I' is fixed and, to use Perry's words, it *exploits* the context to perform its function (i.e., to determine a referent). When we consider how proper names behave in light of this distinction, it is clear that they are not indexicals because, even though Kaplan acknowledges that they have a fixed meaning (or 'character', as Kaplan calls it), their character does not interact with a context to determine a referent (as is the case with indexicals);[50] rather,

[50]The character of a name does not interact with a context because, as Kaplan (1989a:

the context-sensitivity of names is due to the fact that their referents are determined, not by the interaction between a character and a context, but exclusively by contextual factors that operate independently and prior to the deployment of any semantic notions (e.g., the causal history that links an utterance of a proper name in an agent's idiolect to its referent).

Considering that I have argued that names are not indexicals, another strategy to justify extending Kaplan's apparatus to proper names must be found. Here is the suggestion that I endorse. We have seen that Kaplan has doubts with respect to the project of extending his technical apparatus to proper names because the apparatus was developed explicitly to handle context-sensistive expressions such as demonstratives and indexicals, and names are not context-sensitive for him. However, we have also seen that Kaplan acknowledges that the causal history linking the utterance of a name to its referent is a contextual feature. Though Kaplan maintains that this contextual feature of proper names does not warrant treating them as indexicals because it belongs to 'metasemantics'[51] (and the context-sensitivity of indexicals is intimately tied to their character), this contextual feature is still robust enough to warrant extending the distinction between contexts and circumstances of evaluation to sentences involving proper names.

My argument to establish the aformentioned thesis is the following. Re-

562) puts it, 'in the case of proper name words, all three kinds of meaning —referent, content and character— collapse.'

[51]Even though Kaplan (1989b: 573-574) illustrates the distinction between semantics and metasemantics with several examples, he also acknowledges that the distinction is far from clear. However, accepting the distinction is crucial (even if one may disagree as to where it must be drawn) because, without it, one may conflate semantic questions such as 'What is the semantic value of a referring expression r?' and metasemantic questions such as 'How did a referring expression r acquire the semantic value that it has?' Cf. also Stalnaker (1997).

call that Kaplan (1989a: 492) introduces two principles that govern both demonstratives and indexicals:

> Principle 1: The referent of a pure indexical depends on the context and the referent of a demonstrative depends on the associated demonstration.
>
> Principle 2: Indexicals, pure and demonstrative alike, are directly referential expressions.

Shortly after he presents the two principles, Kaplan remarks that they seem to conflict with each other because the first one suggests that the link between an indexical and its referent is a function of something distinct from them and the second one suggests that the link between an indexical and its referent is direct and does not depend on any external factors. To reconcile these two principles, Kaplan (1989a: 494) argues that the distinction between contexts and circumstances of evaluation must be kept in mind and he further maintains that 'once we recognize the obviousness of both principles the distinction between contexts of use and circumstances of evaluation is forced upon us.'

Just as Kaplan introduces these principles (which govern demonstratives and indexicals), one may introduce parallel principles which govern proper names:

> Principle 1*: The referent of a proper name depends on the context of use.
>
> Principle 2*: Proper names are directly referential expressions.

If one is committed (as Kaplan seems to be) to these two parallel principles, the same kind of tension that arises between Principles 1 and 2 also arises between Principles 1* and 2*. And, in light of the fact that Principles 1* and 2* are obvious,[52] the distinction between contexts of use and circumstances of evaluation is forced upon us in the case of names just as clearly as it is in the case of demonstratives and indexicals.

I have argued that Kaplan's distinction between contexts and circumstances of evaluation, which was originally designed to handle demonstratives and indexicals, can be extended to proper names. I want to show now how one can use the distinction to prove that Wreen's argument founders. To do this, it is important first to recall that I mentioned earlier that the contingency of (1) is due to the fact that the causal chain that links in the actual world 'Hesperus' to Hesperus does not exist in several possible scenarios. Since the contingency of (1) is accounted for in terms of the contingency of the causal chain, and this causal chain is, using Kaplan's terminology, a *contextual feature* that enables us to determine the referent of 'Hesperus', one can then hold that (1) is contingent because the referent of 'Hesperus' varies across different *contexts*. This, as I previously indicated, is something that Kripke would readily accept.

However, when Wreen introduces the restriction to a class of possible worlds to derive (8) from (7), he clearly conflates contexts with circumstances

[52] A descriptivist about names might object to the obviousness of Principle 2* by arguing that proper names refer via the mediation of some Fregean sense or individidual concept. However, given that I accepted previously a distinction between two notions of direct reference, the propositional and the Millian, and considered in my subsequent discussion only the propositional notion (see footnote 48 above), I can grant the point raised by the descriptivist and nonetheless still hold on to Principle 2*.

of evaluation. To appreciate this, it is important to keep in mind that one of the main effects of the restriction consists in limiting the possible situations that account for the contingency of (1) to those in which both ends of the causal chain are held fixed but the chain itself does exist. But the very fact that Wreen restricts his attention to these possible situations shows that, when he (1998: 322) writes that 'the possible worlds we're considering are not all the possible worlds', he is no longer treating the possible worlds that he considers as contexts that account for the contingency of (1) but rather as circumstances of evaluation which, after Wreen assumes that the context that generates the proposition expressed by (1) is the actual one, are then mistakenly used to derive the conclusion that 'Hesperus is Phosphorus' is contingent.

To conclude, considering that Wreen violates Kaplan's distinction between contexts and circumstances of evaluation (which, though initially developed for demonstratives and indexicals, can be also extended to names as I have argued) to be able to account for the contingency of (1) and to derive (8) from (7) and that any failure to respect the distinction creates serious difficulties as Kaplan has shown, it is clear that Wreen's argument must be rejected.

3.4 Della Rocca's objection to the necessity of identity

In the previous two sections, I have considered two different challenges to (iii) respectively raised by Gibbard and Wreen and I have shown how these

two challenges can be answered. In the present section, I rehearse an objection that Della Rocca (2002) has raised, not against (iii), but against the necessity of identity. The gist of this objection consists in arguing that, since the proof of the necessity of identity presupposes essentialism and the motivations that Kripke invokes to defend essentialism are subtly self-undermining, the thesis of the necessity of identity cannot be properly established.

To vindicate the necessity of identity from Della Rocca's objection, I proceed in the following way. I present very briefly in §2.4.1 the background discussion in which Della Rocca's objection arises and I outline the objection. Having done that, I argue in §2.4.2 that the necessity of identity can be established without appealing to any essentialist assumptions and I demonstrate that Della Rocca misconstrues Kripke's position concerning the relation between essentialism and the rejection of the use of similarity to account for certain modal claims. Finally, I show in §2.4.3 that there is a good reason that justifies Kripke's discriminatory treatment of certain modal claims that does not involve appealing to the necessity of identity.

3.4.1 Della Rocca's objection

Prior to present Della Rocca's objection, let me rehearse succintly the context in which the objection arises. Della Rocca initially observes that the traditional discussion over whether conceivability is a good guide to possibility or not has important implications because there are several key issues that depend on how it is handled —in particular, the issue of whether certain arguments that have been raised against some physicalist theories of the mind are sound or not. Della Rocca also observes that the overall structure

of these arguments typically involves two claims: a first claim according to which, since it is conceivable that a certain object A has a feature that some object B lacks, it is possible that A has this feature and B lacks it, and a second claim according to which, since there is a modal difference between A and B because it is possible that A has the feature in question and B lacks it, A is distinct from B.

After doing this, Della Rocca brings attention to the fact that, when philosophers have assessed these anti-physicalist arguments, most have concentrated on discussing the first claim (i.e., the inference from conceivability to possibility) and have taken for granted the second claim (i.e., the inference from modal difference to non-identity). Now, instead of joining the debate concerning the first claim, Della Rocca prefers to focus on the second one, which has received very little attention, because he believes that there are considerations that undermine it. Since Della Rocca realizes that he needs to provide some motivation to question a claim that most philosophers take to be obvious, he remarks that the validity of the inference from a modal difference between A and B to a non-identity between A and B presupposes the doctrine that the modal properties of a thing are independent of the way it is referred to (i.e., it presupposes a form of essentialism). Having done that, he then states what his strategy is to put into question the second claim:

> Precisely by starting with the intuitions undergirding essentialism —intuitions that always seemed congenial to me and that concern the role of similarity in modal claims— I will develop a line of argument that shows, I contend, how these intuitions, in a suprising way, undermine themselves or, at least, render them-

selves unjustified. (2002: 225)

Now, just as Della Rocca deploys this strategy to argue against the inference from a modal difference between A and B to a non-identity between A and B, he also relies on the same line of argument to show that the thesis of the necessity of identity cannot be properly established. Indeed, when he considers a proof of the necessity of identity that runs parallel to the one presented in the introduction of this chapter (a proof in which the variables x and y take respectively as substituends the objects a and b), Della Rocca first remarks that this proof presupposes essentialism in order to be valid:

> The inference from the modal properties of a to the modal properties of b goes through only on the assumption that modal properties apply to an object in a reference-independent way. Were we to drop the assumption that modal properties are independent of the way in which the object is picked out, then the above inference would be invalid. Indeed, as far as I can see, any argument for the necessity of identity of a and b crucially presupposes the claim that modal properties are reference-independent. (2002: 227)

After claiming that any proof of the necessity of identity requires presupposing essentialism, Della Rocca points out that, though essentialism has become the orthodox view in these post-Kripkean times, some philosophers remain unconviced by it because it faces two serious problems: the *arbitrariness* worry, which concerns the grounding of modal properties, the *epistemic* worry, concerns our knowledge of modal properties. After drawing attention

to this, Della Rocca then argues that Kripke's strategy to answer these two worries and defend esentialism creates a further problem.

To show this, Della Rocca considers how an anti-essentialist such as Lewis deals with two worries and then compares his strategy with Kripke's. For Lewis, both worries are handled using the notion of *similarity* as the following example illustrates. Granting that Humphrey could have won the 1968 US election, what grounds the modal property that Humphrey has is the fact that there is an individual *similar* in certain relevant respects to Humphrey at some counterfactual situation (i.e., a counterpart of Humphrey) and that this individual is a winner and what accounts for our knowledge that Humphrey has this property is that both the manner and the context in which we refer to Humphrey (which are in principle known to us) emphasize a certain *similarity* relation over many others and this relation determines at some counterfactual situation a counterpart that won the 1968 US election.

On the contrary, Kripke's strategy to deal with the worries consists in dissolving them by arguing that they arise from an erroneous conception of counterfactual situations. Specifically, Kripke argues that there is no real problem as to what grounds the fact Humphrey could have won the election because the demand to ground this fact stems from a picture of counterfactual situations as entities that are given in a purely qualitative way; instead, he suggests that a counterfactual situation at which Humphrey won the election can be determined just by *stipulating* it. These stipulations, however, are neither acts of world-making nor providers of modal knowledge for Kripke, who emphasizes that one can only stipulate what is possible. Ultimately, the source of our modal knowledge, which also acts as a demarcator of the range

of valid stipulations, are our modal intuitions.

After highlighting how Lewis and Kripke respectively handle the two worries, Della Rocca observes that a key element of Kripke's strategy to deal with them (and thus vindicate essentialism) consists in rejecting the use of similarity while upholding stipulation and modal intuition. But this rejection of similarity is not held uniformly by Kripke for all modal claims. Indeed, whereas he (1980: 45-46) emphatically rejects the use of similarity to deal with modal claims such as 'Humphrey could have won the 1968 US election' (a claim Della Rocca calls 'the Humphrey intuition'), he (1980: 104) later acknowledges that, though strictly speaking the modal claim 'Hesperus might not have been Phosphorus' (which Della Rocca calls 'the Hesperus intuition') is false, there is a sense in which the claim expresses something true —namely, the thesis that Hesperus and an object that is qualitatively identical to Phosphorus might be distinct. But the claim has to be interpreted using some notion of similarity to capture adequately this thesis.

It is this difference in the treatment of modal claims that Della Rocca criticizes because Kripke cannot supply seemingly a valid justification for it. Indeed, as Della Rocca (2002: 240) sees things, in order to vindicate essentialism, which is allegedly presupposed by any proof of the necessity of identity, Kripke needs 'to find a relevant difference between the Hesperus intuition and the Humphrey intuition' that explains why the former is to be accounted for in terms of similarity and the latter is not.

For Della Rocca, the only relevant difference between them is that, whereas the Humphrey intuition is consistent with the thesis of the necessity of identity, the Hesperus intuition is not. But Kripke cannot use the inconsistency

between the Hesperus intuition and the necessity of identity to justify reconstruing the former in terms of similarity because, considering that the necessity of identity inevitably requires essentialism, this move amounts to defend essentialism on the basis of essentialism. In light of this, Della Rocca concludes that neither essentialism nor any other thesis that depends on it (e.g., the necessity of identity) can be properly justified.

3.4.2 Essentialism and the necessity of identity

While Della Rocca's objection to the necessity of identity may seem initially quite persuasive, a thorough analysis of his argument reveals the presence of several questionable assertions. One such assertion concerns the role of essentialism with respect to the necessity of identity. Della Rocca claims that anyone who endorses the proof of the necessity of identity given at onset of this chapter must presuppose essentialism in order to justify the validity of the proof. I agree with this. However, he also makes a much stronger claim: that, since *any* proof of the necessity of identity requires presupposing the reference-independence of modal properties, neither Kripke nor anyone else can establish the necessity of identity without begging the question against anti-essentialists. This is precisely where one of my points of contention with Della Rocca lies. In the current subsection, I argue that, since the necessity of identity may be justified *without presupposing* the reference-independence of modal properties, Kripke is not forced to assume essentialism to establish it.[53]

[53]The argumentative strategy that I follow here is exactly the same that Häggqvist (2006: 278-279) deploys, but the argument that I give below is different from Häggqvist's.

Here is an argument for the necessity of identity that does not presuppose the reference-independence of modal properties. Let us assume that there is a distinction, which is often referred to by several medieval authors, between *real relations* and *relations of reason*.[54] According to this distinction, real relations are such that, whenever they obtain between their relata, the relata are changed or modified in a substantive way.[55] To illustrate this, consider the case of two hydrogen atoms that, while standing initially apart, come to be related through a covalent bond. Since the instantiation of the relation alters the reactivity of both atoms by endowing them with a more stable configuration, the relation is clearly a real relation. On the contrary, relations of reason are such that, when they are instantiated by their relata, the relata are not changed or modified. To illustrate this, consider a famous thought experiment devised by Black (1952: 156) that involves supposing that the universe contains nothing but two exactly similar spheres (call them A and B respectively). Since A and B are *two* spheres, there is a relation of numerical distinctness that holds between A and B. And, considering that the instantiation of this relation does not change or alter the relata (which, by assumption, share all their qualities and relational characteristics), it is then patent that numerical distinctness is a relation of reason.

With the distinction in place, one may then argue, as several medieval authors also have done, that identity is a relation of reason. To defend

[54] A detailed discussion of the distinction between real relations and relations of reason in several medieval authors can be found in Brower (2009).

[55] This thesis obviously requires some qualification to specify in which respect the relata of a real relation are modified when it is instantiated. With respect to this, I maintain that what is modified by the instantiation of a real relation are some of the spatio-temporal or causal properties of its relata.

this thesis, consider the following example, which is due to Frege (1892b: 26): if we consider a triangle in which the lines that connect each vertex of the triangle to the midpoint on the opposite side are named a, b and c respectively, the intersection of a and b is identical with the intersection of b and c. As we can appreciate, the identity relation is clearly instantiated in this case but its instantiation does not modify or change the geometrical properties of the relata —or, speaking more accurately, of the object that is picked by both relata since identity holds between any object and itself.[56]

Granting that identity is a relation of reason, the next step in my argument consists in defending the thesis that, if a relation of reason is instantiated, it is then necessarily instantiated. This thesis can be supported using the following reasoning. Some authors have argued that the causal and spatio-temporal structure of the world is metaphysically contingent in the sense that the world could have been such that no causal or spatio-temporal relations were instantiated at it.[57] However, even if the world had turned out to be such that no causal or spatio-temporal relations were instantiated at it, relations of reason would still be instantiated by whatever objects happened to exist at that world. For instance, if only sets and other mathematical entities (which are presumed to be necessary existents) had existed, it would have still been the case that the relation of numerical distinctness is instan-

[56]There is a persistent perplexity about this issue that is often expressed through the following question: if identity relates an object to itself, what is relational about it? Regarding this perplexity, I adopt Quine's answer (1960: 116) which consists in tracing it to a confusion between object and sign: "what makes identity a relation, and '=' a relative term, is that '=' goes between distinct occurrences of singular terms, same or distinct, and not that it relates distinct objects."

[57]For instance, Lowe (1993) has suggested that the world could have been such that only abstract objects existed.

tiated between, say, $\{\{\emptyset\}\}$ and $\{\emptyset, \{\emptyset\}\}$. In light of this, it seems that, if relations of reason are instantiated, they are necessarily instantiated.

Consequently, granting the distinction between real relations and relations of reason, we can give a proof of the necessity of identity by using modus ponens on the following premises: relations of reason are necessarily instantiated if they are instatiated at all and identity is a relation of reason. With regard to this argument, I acknowledge that it depends on some controversial assumptions (in particular, it presupposes a Platonist conception of mathematical entities in order to substantiate the conditional premise). But the important point is that it does not presuppose the reference-independence of modal properties. Thus, *pace* Della Rocca, Kripke is not forced to assume essentialism to establish the necessity of identity.

Before moving on to the next subsection, I want to bring attention to another issue. Della Rocca (2002: 234) claims that, for Kripke, the rejection of the use of similarity is necessary for 'defending essentialism against the kind of similarity account of modal claims that Lewis and other anti-esentialists offer.' However, this betrays a serious misunderstanding of Kripke. For Kripke, the rejection of the use of similarity to account for modal claims does not function as a tool to defend essentialism; rather, it is a *consequence* of essentialism. To be more specific, it stems from the adoption by Kripke of a particular conception of counterfactual situations. While Lewis maintains that counterfactual situations are given in a purely qualitative way and other authors (e.g., Carnap 1947) hold that counterfactual situations are linguistic constructs that provide descriptions of total states of the world, Kripke (1963) endorses a view of counterfactual situations as 'points' the internal

structure of which is given as primitive. It is this assumption concerning the nature of counterfactual situations —which is an essentialist assumption— that justifies the rejection of the use of similarity to account for modal claims, and not the other way around.

Putting this in other terms, while Kripke rejects the use of similarity to account for certain modal claims, this rejection is not used to uphold essentialism: it is derived from essentialism, which is built into the inner structure of counterfactual situations. The vindication of essentialism against similarity-based accounts of modal claims is undertaken by Kripke in a different way: he argues that these accounts are undermined by serious formal difficulties.[58] Thus, even though Della Rocca is right in pointing out that Kripke is engaged in a fierce debate with anti-essentialists over the proper way of interpreting modal claims, his view of the dialectical situation is mistaken because Kripke does not defend essentialism by rejecting similarity-based accounts of modal claims. Rather, Kripke's defense of essentialism is grounded primarily on formal considerations, and one of the consequences of this defense is the rejection of the use of similarity for certain modal claims.

[58]In particular, Kripke (1980: 45fn13) complains that the principle $(y)[(x)\mathrm{A}(x)\supset\mathrm{A}(y)]$ fails on these accounts if $\mathrm{A}(x)$ contains modal operators. Lewis (1983b: 45-46) argues in a rejoinder that this principle and others only fail if we do not recognize that modal operators conceal quantifiers that bind variables. Although I disagree with the rejoinder, a conclusive refutation is beyond the scope of the present work because it would involve a systematic analysis of the principles of modal logic to assess if they are validated by counterpart theory. Instead, I want to suggest here a strategy that can be used to *doubt* Lewis' rejoinder.

Since Lewis assumes that modal operators conceal quantifiers, he interprets modal notions in quantificational terms. However, there are alternative accounts of modal notions; for instance, some authors construe them in terms of primitive *modes* that modify the status of instantiation of properties. For some reasons to prefer this interpretation over the quantificational account, see McGinn (1981) and (2001).

3.4.3 A justification of Kripke's discriminatory treament of modal intuitions

I turn now to another point of contention. Della Rocca rightly observes that Kripke does not advocate a uniform approach to modal claims: some (e.g., the Humphrey intuition) are not to be interpreted using similarity relations while others (e.g., the Hesperus intuition) need to be construed in terms of similarity. While this assymetric approach seems initially plausible, Della Rocca argues that Kripke cannot provide ultimately a good reason for it.

In this subsection, I argue that there is a good reason to justify the assymetric treatment that the Humphrey intuition and the Hesperus intuition receive, i.e., a reason that explains why the Hesperus intuition is to be treated in terms of similarity (and the Humphrey intuition is not) without invoking the necessity of identity. Before presenting my suggestion, however, let me introduce a proposal due to Bealer (2004). Indeed, although I believe that Bealer's proposal ultimately does not provide such a reason, its failure is of great interest because it points towards what I take to be a satisfactory solution.

To understand Bealer's proposal, let me sketch briefly the main lines of his paper. Bealer wants to show that, in spite of the fallibility of modal intuitions, their evidential status is not under threat because erroneous modal intuitions can in principle be identified and dealt with by subjecting all modal intuitions to a certain a priori dialectic. The strategy that he adopts to argue for this thesis involves two basic phases: he initially establishes a taxonomy of the traditional sources of erroneous modal intuitions (which he calls 'modal

errors') and, having done this, he argues that, in each case, there is a way to recognize and handle appropriately these modal errors.

In what follows, I focus exclusively on Bealer's discussion of one particular source of modal errors (namely, the conflation between metaphysical possibility and epistemic possibility). I do this because, after Bealer criticizes what he takes to be Kripke's way to handle this source of modal errors and presents his own proposal, he also suggests a way in which the challenge that Della Rocca raises against Kripke may be answered. To appreciate more clearly Bealer's reasoning, I rehearse in the first place how Kripke handles certain erroneous modal intutions (e.g., the Hesperus intuition) according to Bealer; I turn subsequently to present Bealer's criticism of Kripke and characterize his counterproposal and, finally, I show why Bealer's counterproposal is problematic.

It is well-known that, while Kripke holds that it could not be the case that Hesperus is distinct from Phosphorus, he (1980: 104) also grants that 'it could have turned out that Hesperus was not Phosphorus.' Bealer maintains that, to reconcile these two claims, Kripke argues that both do correspond to genuine modal intuitions that do not really conflict; what generates the illusion of inconsistency is that one of them is *misreported* and the reports conflict with each other. Bealer also holds that the solution that Kripke suggests to correct the misreport and eliminate the illusion consists in interpreting the second claim as really describing a counterfactual situation in which Hesperus and an object that is qualitatively identical to Phosphorus are not identical.

However, Bealer does not agree with Kripke's view that the source of

modal errors lies in a conflict between the reports of the modal intuitions because it has problematic consequences. Indeed, he (2004: 29) argues that this view calls for a meta-linguistic rephrasing strategy in which statements of the form 'It could have turned out that P' have to be analyzed as stating "It is possible that a population of speakers in a situation qualitatively identical to ours would make a true statement by uttering 'P' with normal literal intent", and he highlights a series of difficulties that undermine this strategy.[59]

As an alternative, Bealer suggests denying that there is even a conflict between the reports of modal intuitions and maintaining that the source of modal error lies in a failure to distinguish that 'could' has different senses in 'It could not be the case that Hesperus is distinct from Phosphorus' and in 'It could turn out that Hesperus was not Phosphorus'. And he also suggests that, once we realize that 'could' has at least two different senses that correspond to metaphysical and epistemic possibilities respectively, the modal error dissolves.

Though Bealer's suggestion is *prima facie* attractive for somebody who has both modal intuitions and attempts to show that they do not conflict by arguing that they correspond to distinct types of possibility, he remarks that his proposal generates a doubt: how can we determine if a modal intuition corresponds to a metaphysical possibility or to an epistemic possibility? The following passage captures his answer to the question:

> Suppose you are considering the modal status of one of the classi-

[59]In particular, Bealer points out that this rephrasing strategy violates the Langford-Church translation test according to which, to assess whether an analysis is correct, one has to translate the analysans and the analysandum into another language and then evaluate if the translations really convey the same meaning.

> cal hypothetical-case propositions P (e.g., that Aristotle was not the teacher of Alexander) (...) Typically, for the philosophical purpose at hand, you may bypass the question of whether P is metaphyscially possible and consider instead whether P is a contingent proposition. Therefore, one may recast without loss our question in that idiom. And since a survey of cases shows that 'contingent' —unlike 'possible', 'could' and so forth— does not have an epistemic reading, one need no longer worry about the possibility of equivocation. (2004: 30-31)

Unfortunately, the proposal that Bealer makes here only works if we accept that 'contingent' admits exclusively a metaphysical reading, and this thesis is rejected by some authors. For instance, Sidelle (1989, 2002) has argued that, while there is a sense in which the sentence 'Hesperus is Phosphorus' is necessary, there is also a sense in which it is contingent.[60] 'Hesperus is Phosphorus' is *linguistically* necessary according to Sidelle in virtue of the existence of a certain analytic principle of the form "Nothing counts as Hesperus unless it has same deep explanatory features as the object we call 'Hesperus'." And it is *epistemically* contingent because it is established empirically that the object that has the deep explanatory features of the object we call 'Hesperus' is Phosphorus. Thus, since the method of rephrasing our modal intuitions in terms of the notion of contingency to evaluate if they correspond to metaphysical or epistemic possibilities is only effective if we make certain assumptions (and these assumptions are disputed by conventionalists

[60]Sidelle discusses the sentence 'Water is H_2O' rather than 'Hesperus is Phosphorus', but the same observations that he makes concerning the first sentence can be carried over to the second one.

about necessity such as Sidelle), the use of Bealer's rephrasing strategy can be objected by arguing that it begs the question against conventionalism.

Although I believe that conventionalism about necessity is mistaken, relying on a method that begs the question against it to answer Della Rocca's challenge is not appealing. Accordingly, let me introduce an alternative suggestion to address Della Rocca's challenge. To do this, it is important to notice first that Bealer's characterization of Kripke's views on modal error is mistaken. Indeed, Kripke does not maintain, when he considers both the intuition that it could not be the case that Hesperus is distinct from Phophorus and the intuition that it could turn out that Hesperus was not Phosphorus, that one of the intuitions is misreported. That is not the source of modal error for Kripke. The real source comes to light when we consider closely the following passage in which Kripke discusses the use of similarity for certain modal intuitions:

> *This* table itself could not have had a different origin from the one it in fact had, but in a situation qualitatively identical to this one with respect to all the evidence that I had in advance, the room could have contained *a table made of ice* in place of this one. Something like counterpart theory is thus applicable to the situation, but it applies only because we are *not* interested in what might have been true of *this particular* table but in what might or might not be true of *a table* given certain evidence. (1980: 142)

As the second sentence of the passage illustrates, the source of modal error lies not in the fact that a certain modal intuition (e.g., the modal

intuition that the table could turn out to be made of ice) is misreported, as Bealer maintains, but rather in the fact that, though we often appeal to our modal intuitions to evaluate the modal profile of an object, we tend to conflate our modal intuitions about the object with our modal intuitions about some of its properties.[61] For Kripke, it is only when we are concerned with certain properties of the object rather than with the object itself that an interpretation of our modal intuitions in terms of similarity relations is legitimate. In other cases, when we are primarily concerned with the object (i.e., when we are interested in the *de re* genuine modalities), interpreting our modal intuitions in terms of similarity is, as Kripke puts it, perverse.

In light of the previous discussion concerning the real source of modal error for Kripke, we can provide an explanation of why the Humphrey intuition and the Hesperus intuition have to be treated differently without invoking the necessity of identity.[62] When we consider the Humphrey intuition as Kripke initially introduces it, we can appreciate that it is an intuition about an object (namely, Humphrey) and not an intuition about some properties that Humphrey has such as his looks or his political views:

> If we have such an intuition about the possibility of *that* (*this man's* electoral loss), then it is then it is about the possibility of that. It need not be identified with the possibility of a man looking like such and such, or holding such an such political views,

[61] In a previous paper, Bealer (2002: 83) briefly considers this approach, but he does not pursue it.

[62] Though I am content with showing that there is a reason that justifies Kripke's discriminatory treatment of the Humphrey intuition and the Hesperus intuition, I believe the account that I present can be generalized to other cases. I intend to pursue this project in some future work.

> or otherwise qualitatively described, having lost. We can point to the *man* and ask what might have happened to *him*, had things been different. (1980: 46)

Consequently, since the Humphrey intuition concerns the counterfactual on-goings of a certain *object*, it would be a serious mistake to interpret it in terms of similarity. On the contrary, when we consider the Hesperus intuition, we can appreciate that, although the intuition seems to be initially about an object (i.e., Hesperus), it really turns out to be about certain *properties* of that object. Indeed, after Kripke (1980: 103) grants that 'in advance we are inclined to say that the question of whether Hesperus is Phosphorus might have turned out either way', he points out that, when we try to determine which truth (if any) is expressed by this modal intuition, the truth in question concerns, not Hesperus, but an object that has the same visible properties as Hesperus (e.g., properties such as *occupying the same position that Hesperus has in the evening*). It is precisely because of this that the Hesperus intuition has to be reconstrued in terms of similarity. To conclude, there is a way to justify Kripke's assymetric treatment of the Humphrey intuition and the Hesperus intuition which consists in arguing that, since the Humphrey intuition concerns an object whereas the Hesperus intuition concerns (notwithstanding the initial appearaces) certain properties of an object, only the Hesperus intuition requires being interpreted in terms of similarity.

3.5 Conclusion

It is now time to recap. In the previous sections, I have argued that some of the main arguments developed to challenge both the thesis that true identity statements between names are necessary and the necessity of identity are flawed on inconclusive. In particular, I have argued that Gibbard's arguments based on the Goliath-Lumpl scenario all founder. I have also shown that the objections raised by Wreen and Della Rocca can be effectively dealt with. If my defense of the necessity of identity has been indeed succesful, my efforts here can be seen as a partial vindication of other arguments that rely on the necessity of identity (e.g., Kripke's modal argument for property dualism) as well as a partial defense of the classical picture of identity.

Chapter 4

The determinacy of identity

4.1 Introduction

There are many things usually taken to display indeterminacy. For instance, some argue that mountains and other geographical features (rivers, forests, cities, oceans, etc.) display indeterminacy because there is no fact of the matter as to where their borders lie.[1] Others claim that fundamental particles such as electrons exhibit indeterminacy because, although they can be conceived as individuals exhibiting features such as location or momentum, there are circumstances in which there is no fact of the matter as to whether they exhibit a certain location or momentum.[2] Moreover, there are also others who hold that some of our fundamental concepts such as that of existence exhibit indeterminacy because, if we suppose that the activity

[1] See, for instance, the papers collected in Burrough and Frank (1996). For a criticism of this view, see Varzi (2001).

[2] For instance, see Maudlin (2003: 470-472). For alternative ways of defending the thesis that fundamental particles exhibit indeterminacy, see French and Krause (2003).

of certain objects constitutes the life of a certain biological organism (say, a cat) and it is indeterminate whether a group of objects exhibit the relevant kind of activity, it is then indeterminate whether a cat exists or not.[3]

In light of the abovementioned views, the following questions arise: considering that many things are deemed to exhibit indeterminacy, is *identity* among them in some circumstances? Is identity indeterminate in certain cases?[4] *Prima facie*, there seem to be several puzzles and thought experiments that provide support for an affirmative answer for both questions. For instance, Parsons (2000: 2-4) claims that the introduction of some variations to the Ship of Theseus story provides a strong motivation to maintain that identity is indeterminate in certain cases.

To appreciate this, let us consider first the unmodified story. Let us suppose that, after a ship leaves its port (call the ship at this time 'Alpha'), its old wooden planks are slowly removed one by one and replaced by new planks until none remains (call the ship at this time 'Beta') and suppose further that the old wooden planks are saved by someone who reassembles them into a ship (call this ship 'Gamma'). While it is unobjectionable that Beta is distinct from Gamma, Parsons remarks that philosophers are often split on the question of whether Alpha is identical to Beta or Alpha is identical to Gamma: whereas some defend the thesis that Alpha and Beta are identical on the basis of considerations about spatio-temporal continuity, others defend

[3]For instance, see Van Inwagen (1990), chap. 18 and Hawley (2002).

[4]It is important to notice that either an affirmative or a negative answer to either one of these two questions has wide repercussions in virtue of the foundational role that the identity relation plays in our conceptual scheme. For instance, if we maintained that identity is indeed indeterminate in certain cases, we would be forced to admit that notions such as synonymy or simultaneity, which are grounded on the concept of identity, also exhibit indeterminacy in certain cases.

the thesis that Alpha and Gamma are identical on the basis of considerations about material constitution.[5]

However, the introduction of small variations to the scenario is sufficient to generate doubts about these two alternatives. As Parsons notes, if we suppose that the wooden planks of Alpha are substituted by planks made of a material completely different from wood, the plausibility of the thesis that Alpha is identical to Gamma is severly weakened. Similarly, if we suppose that, instead of using all the planks of Alpha to assemble Beta, we use only some of the planks of Alpha while the rest of planks are new, the plausibility of the thesis that Alpha is identical to Beta is also severely weakened.

The moral that Parsons draws from this puzzle is that, since every attempt to defend either the identity of Alpha and Beta or the identity of Alpha and Gamma can be ultimately undermined by tinkering slightly the story in a way that smudges our intuitions, these intuitions have to be rejected. For Parsons (2000: 5), this rejection is ultimately achieved by accepting that 'there is no fact of the matter regarding whether [Gamma] is [Alpha] or whether [Beta] is [Alpha].'

In addition to Parsons, other authors also present puzzles and thought experiments that support the thesis that identity is indeterminate. Consider the following thought experiment, which is presented by Broome (1984: 6-7). Suppose that a certain club is formed by a group of like-minded persons,

[5]Parsons observes that the spectrum is not exhausted by these two options. In addition, some philosophers reject that Alpha is identical to Beta because they endorse mereological essentialism (i.e., the thesis that all the parts of an object are essential to it) and the thesis that Beta because Alpha is identical to Gamma because they embrace the thesis that identity requires not only spatio-temporal continuity but continuity of shape and function. Parsons leaves aside these views on the grounds that they are too extreme and have very few supporters.

holds meetings for a period but later supends them as most of its members are no longer able to attend them. Suppose further that, after a long period of time, some of the original members start to meet again using the same name. If we grant, as Broome persuasively suggests, that a club is constituted by its rules and society's conventions and if we also suppose that no rules or provisions exist concerning what counts as the demise of the club, we seem to have a case of indeterminate identity because there appears to be no fact of the matter regarding whether the club prior to the interruption of the meetings is identical to the club after the resumption of the meetings.

Another thought experiment due to Van Inwagen (1990: 241-242) also motivates the thesis that identity is indeterminate. Suppose that a certain person (call her Alice) enters an infernal machine called the Cabinet that performs on her an experiment such that it is indeterminate whether the activity of the objects that constitute Alice's life is terminated or not by the experiment.[6] After the experiment, a person (call her Celia) emerges from the cabinet. Considering that, according to Van Inwagen, Alice and Celia are identical if their lives are one and the same, if Alice enters the Cabinet and Celia emerges subsequently from it, it seems that there is no fact of the matter concerning whether Alice is identical to Celia or not.

Though the aformentioned puzzles and thought experiments are extremely convincing, none of them are real arguments that support the thesis that

[6]Given that different philosophers hold different views on what kinds of events would cause the termination of Alice's life (e.g., some maintain that the termination of Alice's life requires something that permanently disrupts her psychological continuity while others hold that the termination of Alice's requires an event that destroys her body), Van Inwagen supposes that the Cabinet is such that it can be set to perform different kinds of experiments on Alice in order to generate a situation in which our intuitions, whatever they are, dictate that it is indeterminate whether Alice's life is terminated or not.

identity is indeterminate. In light of this, although the thesis seems correct *prima facie*, it lacks the kind of firm support that is provided by genuine philosophical arguments. But the situation is far worse for supporters of the thesis than the mere lack of an argument for it suggests: some philosophers, in particular Evans (1978) and Salmon (2005b), have developed arguments that prove that the thesis is incoherent because, from certain plausible assumptions, one can derive from it an absurdity.

Both arguments have exerted a tremendous influence on the debate concerning whether the thesis is true or not, with several philosophers taking position against them to vindicate the intuitions underpinning the abovementioned puzzles and thought experiments and other philosophers defending the arguments from objections raised by members of the first group. However, notwithstanding all the attention that they have received, some serious discrepancies concerning the proper interpretation of the arguments still linger. In particular, in the case of Evans' argument, philosophers have expressed radically distinct views about which specific thesis Evans aims to refute using his argument, and this has had an impact on the attitudes that some have adopted with respect to it.

The main goal of the present chapter is to argue that the thesis that identity is indeterminate under certain circumstances is, in a sense to be specified further down, false. Thus, I maintain that, for any objects a and b, if a is identical to b, then it is determinate whether a is identical to b. The general strategy that I pursue to prove this conclusion consists in organizing a systematic defense of Evans' argument by addressing a series of misunderstandings about it as well as several objections that have been raised against

it from several places.[7]

Here is a roadmap that details how I proceed below. In §4.2, I present briefly Evans' argument after rehearsing the background considerations that, according to Evans, help generate the thesis that he aims to disprove in his paper. Having done that, I turn to some interpretive matters in §4.3. In particular, I distinguish two ways to interpret the thesis that identity is indeterminate in certain cases, which correspond to two distinct views about the source of indeterminacy. After clarifying what these two views consist in, I show that they correspond to two distinct interpretations of Evans' argument, one of which I take to be the one that Evans intended on the basis of certain remarks made by Lewis.

Once these interpetive matters are settled, I consider several objections that have been raised against the argument. In §4.4, I introduce and critically discuss two objections raised by Parsons and Woodruff (1995). The gist of the first objection is that Evans assumes without argument that a lambda abstract that binds a variable governed by the indeterminacy operator '$\bigtriangledown$' both expresses a property and fully satisfies the principle of lambda abstraction. The second objection is that, although Evans invokes Leibniz's Law in order to derive from his initial assumption a thesis that contradicts it, what he really employs is the contrapostitive of Leibniz's Law and, given that the contrapositives of some valid principles and arguments are invalid in a logic that sanctions indeterminacy of truth-value, he begs a question of logic when he assumes that the contrapositive of Leibniz's Law is valid. After rehearsing

[7]I focus here on Evans' argument rather than Salmon's because, in virtue of its formulation, Evans' argument has been frequently the object of several misapprehensions, which I aim to dispel. Salmon's argument, which is more sober and elegant, has elicited less controversy.

both objections in more detail, I show that both can be met on the basis of certain considerations stemming from the nature and the role of properties and from the interpretation of Evans' argument in a non-classical logic that sanctions indeterminacy of truth-value.

Subsequently, I examine in §4.5 a challenge raised against Evans' argument by Van Inwagen (1990). Van Inwagen's challenge consists in arguing that some of the inferences in the argument are invalid under a plausible construal. His general strategy to do this consists in developing a semantics for a fragment of the language of first-order logic that includes the key elements used in Evans' argument —namely, identity, lambda abstraction and the indeterminacy operator '$\bigtriangledown$'. Once the semantics is in place, Van Inwagen shows that there are certain interpretations of that language fragment in which some of the inferences of Evans'argument clearly fail to be truth-preserving. On this basis, Van Inwagen then concludes that there are cases in which identity is indeterminate. In response to Van Inwagen, I show that the semantics he proposes is undermined by serious difficulties.

In §4.6, I offer a critical appraisal of a response to the argument put forth recently by Williams (2008). According to Williams, it is possible for an advocate of the thesis that identity is indeterminate in certain cases to resist Evans' argument by exploiting a certain loophole in his reasoning. Specifically, Williams argues that, if one can establish that there are certain scenarios or cases in which referential indeterminacy arises from worldly indeterminacy, the argument can be rejected on the grounds that, on one of its interpretations (namely, the one that I embrace as correct in §4.3), the existence of these scenarios undermines the conclusion that Evans argues for.

After highlighting the loophole, Williams exploits it in the following way: he presents us with a certain framework in which worldly indeterminacy arises and then, after doing this, he shows, using a scenario in which an amoeba undergoes fission, how this scenario gives rise to a case of indeterminate identity within the framework. The gist of my reply to Williams consists in showing that, although there are indeed cases in which worldly indeterminacy produces referential indeterminacy, the *kind* of worldly indeterminacy that Williams considers is unable to engender the *kind* of referential indeterminacy that he needs to exploit the loophole and challenge Evans' argument.

I consider in §4.7 another recent response to Evans' argument formulated by Barnes (2009). In a nutshell, Barnes argues that supporters of indeterminate identity may refuse to accept the conclusion of the argument because, if the indeterminacy operator '$\triangledown$' that Evans uses is interpreted in counterpart-theoretic terms, some inferences of the argument fail to be valid all together. She also argues that the counterpart-theoretic interpretation of '$\triangledown$', which she presents within a framework in which indeterminacy is treated as a modal notion, is strongly motivated not only by certain remarks of Evans himself but also by some considerations stemming from (a) certain similarities between the semantic apparatus deployed by counterpart theorists to interpret modal claims and the semantic apparatus she adopts to handle indeterminate statements and (b) the context-sensitivity of indeterminate statements. To vindicate Evans' argument, I show that the motivations that Barnes puts forth to interpret '$\triangledown$' in counterpart-theoretic terms can be succesfully countered by invoking certain remarks that Barnes herself makes about the semantic apparatus she embraces for indeterminate statements and

by considering alternative ways to handle the context-sensitivity of indeterminate statements. Finally, I offer in §4.8 a brief conclusion.

4.2 Evans' argument

Before presenting the argument, it is important to have a clear understanding of the assumptions from which stems the thesis that Evans aims to refute in his paper. Indeed, this understanding will allow us to ward off certain mistakes and confusions that are often made with respect to the argument. Evans (1978: 208) begins his paper by noting that, whereas some people claim that 'the world itself might be vague [in the sense that] rather than vagueness being a deficiency in our mode of describing the world, it would then be a necessary feature of any true description of it', others maintain that 'amongst the statements which may not have a determinate truth-value as a result of their vagueness are identity statements.' These two views can be more clearly formulated as follows:

> (WI) It might be the case that the only indeterminacy there is is worldly indeterminacy.
>
> (IS) Identity statements are among the statements that may be indeterminate in truth-value.

After presenting them, Evans (*ibid.*) argues that by 'combining these two views we would arrive at the idea that the world might contain certain objects about which it is a fact that they have fuzzy boundaries.' A possible interpretation of the idea that Evans expresses in this passage is captured by the following thesis:

(FB) There might be certain objects with fuzzy boundaries and, if these objects figure in identity statements, the statements might be indeterminate in truth-value.

However, if we interpret the abovementioned passage as expressing FB, this seems to be a rather narrow interpretation. Indeed, one might concede that, if the way in which worldly indeterminacy is manifested is through the existence of objects with fuzzy boundaries and these objects with fuzzy boundaries are involved in identity statements, these identity statements may be indeterminate in truth-value.[8] But aren't there other manifestations of worldly indeterminacy besides fuzziness in boundaries? Some authors, as I mentioned in §3.1, hold that existence exhibits indeterminacy. Others (e.g., Barnes 2005) suggest that the exemplification of sparse properties (i.e., the basic, irreducible properties that constitute the bedrock of reality) also displays indeterminacy. In light of this, I believe that a broader and also more interesting interpretation of the idea that combines both WI and IS is expressed by this thesis:

(IT) Some identity statements are of indeterminate truth-value exclusively in virtue of worldly indeterminacy (in any of its manifestations).

[8] To illustrate this, consider the following example, which is introduced by Lewis (1993): let us suppose that Tibbles the cat is shedding and so there is no fact of the matter concerning where his boundaries lie because some of the hairs are neither definitely parts of Tibbles nor definitely not. In addition to Tibbles, there are several cat-like entities that only differ from him by lacking one of the aforementioned hairs. Call one of these entities Tibbles_1. If we are asked whether Tibbles is identical to Tibbles_1, a *prima facie* tempting response to the question is that it is indeterminate whether Tibbles is Tibbles_1 because it is indeterminate whether the fuzzy boundaries of Tibbles match those of Tibbles_1.

Even though I take IT to be the primary target of Evans' argument, I want to postpone until §4.3 my reasons for endorsing this interpretation over others. Grant me, for the time being, that IT is the thesis that Evans aims to undermine and let me lay down the argument, which goes as follows. Evans first supposes that 'a' and 'b' are two singular terms such that the identity statment 'a=b' is indeterminate in truth-value and he develops a formal notation to express this thesis through the introduction of the sentential operator '$\triangledown$', which is read as 'It is indeterminate whether...'. Thus, the initial supposition can be thus expressed:

(1) $\triangledown(a{=}b)$.

Evans argues that this supposition expresses a fact about b —namely, the fact that it has the property expressed by the lambda abstract '$\lambda x[\triangledown(x{=}a)]$'. Thus, we have from (1) by property abstraction (or, as it also called, lambda abstraction):

(2) $\lambda x[\triangledown(x{=}a)]$b.

Now, given that it is a principle of logic that every object is determinately identical to itself, we have:

(3) $\neg\triangledown(a{=}a)$.

Just as (2) expresses a fact about b, Evans claims that (3) expresses a fact about a —namely, the fact that it lacks the property expressed by the lambda abstract '$\lambda x[\triangledown(x{=}a)]$'. Thus, we have from (3) by the contrapositive of property concretion (which states, for any property F and any object o, that if $\neg Fo$ then o lacks F):

(4) $\neg\lambda x[\bigtriangledown(x{=}a)]a$.

And, given that (2) and (4) respectively state that b has the property expressed by the lambda abstract and that a lacks it, we can then conclude by applying the contrapositive of Leibniz's Law to them that:

(5) $\neg(a{=}b)$.

Although (5) is not the negation of the initial supposition, Evans nonetheless maintains that (5) contradicts (1) in the sense that what is originally assumed by the argument is that the identity statement '$a{=}b$' is indeterminate in truth-value. Now, if '$a{=}b$' is indeterminate in truth-value, so is presumably '$\neg(a{=}b)$' but this is at odds with (5), which asserts that it is true that a is distinct from b. In light of this, the basic structure of the argument is that of a *reductio* which shows that the very thesis of indeterminate identity is incoherent because any supposed case of indeterminate identity turns out to be a case of non-identity.

Before turning to interpretive matters in the next section, I want to highlight a couple of issues about the structure and the formal machinery that the argument relies on, since these issues have an impact on certain responses that have been addressed to the argument. The first issue, which concerns the structure of the argument, is the following: given that Evans considers the argument as a *reductio* on the grounds that (5) contradicts (1), he does not need to go any further than (5) to make his case. Putting this in slightly different terms, even though Evans could appeal to a variety of principles in order to derive from (5) the straightforward negation of (1), he is not required to do so to establish his point. Consequently, even though Evans

suggests in the last paragraph of his paper a method to obtain a straightforward negation of (1),[9] he does not really embrace it, as Heck (1998: 294-295) has persuasively argued. In light of this, some objections to Evans' argument which hinge on showing that the suggestion he makes at the end of his paper is problematic can be safely dismissed on the account that they are concerned with what is really an aside.[10]

The second issue I want to highlight is the paucity of the formal machinery used in Evans' argument. In addition to the supposition for *reductio*, only three principles (i.e., property abstraction, contrapositive of property concretion and contrapositive of Leibniz's Law) and one assumption (i.e., the non-indeterminacy of self-identity) are needed to derive the contradiction. The language that Evans employs to formulate the argument also has very few elements: some individual constants ('a' and 'b'), a variable ('x'), the negation operator '$\neg$', lambda expressions ('$\lambda x[...x...]$'), the identity sign '$=$' and a sentential indeterminacy operator '$\bigtriangledown$' that is used to express indeterminacy in truth-value. No other principles or elements are invoked by Evans in the proof. In virtue of this, it is plausible to maintain that Evans does not require to assume any special 'logic of indeterminacy' to derive the contradiction. Indeed, as far as I see, neither the principles that Evans relies on nor the language that he employs call for anything beyond classical logic. Thus, although the argument may be interpreted (like any other argument) within a framework distinct from that of classical logic, this is not

[9]Here is Evans' suggestion (*ibid.*): "If 'Indefinitely' and its dual 'Definitely' ('$\bigtriangleup$') generate a modal logic as strong as S5, (1)-(4) and, presumably, Leibniz's Law may be strengthened with a 'Definitely' prefix, enabling us to derive: (5') $\bigtriangleup\neg(a{=}b)$"

[10]A point in case is an objection raised by Over (1989) who argues that, if '$\bigtriangledown$' and '$\bigtriangleup$' are indeed duals, we may derive from (3) the thesis that a is definitely not identical to a.

presumably something that Evans intended.

4.3 Some interpretive matters

Having presented Evans' argument as well as underscored certain issues that stem from its structure and the formal machinery it employs, we can turn now to some interpretive matters. As I mentioned in §4.1, the thesis that many authors appear to support through several puzzles and thought experiments and that Evans intends to refute is that identity can be indeterminate in certain cases. But how should we understand this thesis? A thorough reflection on the puzzles and thought experiements previously introduced shows that there is a view that surfaces in all of them —namely, the view that in certain cases the identity between an object a and an object b is indeterminate because there is no fact of the matter concerning whether the identity statement 'a=b' is true or not. Let us call this view, which follows from IS, the Indeterminate Identity Thesis:

> (IIT) There are certain identity statements that may be of indeterminate truth-value.

According to IIT, there are cases in which there is just no fact of the matter regarding whether an identity statement 'a=b' is true or not. IIT provides us with a plausible way to understand the initial thesis that identity is indeterminate in certain cases, but it still leaves certain questions unanswered. Indeed, even though one may accept that some identity statements may be of indeterminate truth-value, this does not indicate to us the *source* of their indeterminacy. What makes it the case that these identity statements are

such that there is no fact of the matter regarding whether they are true or not?

There are two contrasting views on this issue. The first view holds that, for all identity statements that exhibit indeterminacy of truth-value, this indeterminacy stems exclusively from *semantic indecision*. To illustrate what the view consists in, consider first the following classical example of referential indeterminacy stemming from semantic indecision, which is due to Quine (1960, chap. 2): the term 'gavagai', which is uttered by a native in front a linguist when the native receives a series of stimuli that correspond to those generated by the presence of a rabbit, could in principle be used to refer to a rabbit, or to an undetached rabbit-part or to a temporal rabbit-stage (since all those entities are also present when a rabbit is), but there is no way to determine which candidate is the correct one. Granting that none of these potential candidates for being the reference of 'gavagai' are indeterminate entities (e.g., objects with fuzzy boundaries), the identity statement "The referent of 'gavagai'=The rabbit responsible for the stimuli perceived by the native" is then indeterminate in truth-value exclusively because there is no fact of the matter regarding whether 'gavagai' refers to the precise rabbit in question or not. Now, the view, which holds that all indeterminate identity sentences are so for the same basic reason that the aforementioned sentence is, is captured by the following thesis:

> (IIT_s) If an identity statement is indeterminate, it is so exclusively in virtue of semantic indecision.

The second view holds that, in the case of some indeterminate identity statements, their indeterminacy in truth-value stems, not from the fact that

at least one of the *terms* in the identity statements in question exhibits semantic indecision, but rather to the fact that the *objects* referred to by the terms in the identity statements are indeterminate. It is important to notice that this view does not preclude the existence of indeterminate identity statements which are so in virtue of semantic indecision: it merely affirms that, among indeterminate identity statements, the indeterminacy in truth-value of some of them arises from some indeterminacy in the *entities* themselves rather than in the *words* we use to pick them out.

An illustration of this view is provided by the fate of the unfortunate person who steps into Van Inwagen's Cabinet. Indeed, even though it is determinate that 'Alice' refers to the person that steps into Cabinet and that 'Celia' refers to the person that steps out of the Cabinet, the identity statement 'Alice=Celia' is indeterminate in truth-value and, since its indeterminacy cannot be traced to any semantic indecision, the only way to explain it is, according to Van Inwagen, by tracing it to some form of *worldly* indeterminacy (namely, the indeterminacy regarding whether Alice continues to exist or not after stepping into the Cabinet). This view is neatly expressed by the following thesis:

> (IIT_w) Some identity statements are of indeterminate truth-value exclusively in virtue of worldly indeterminacy (in any of its manifestations).

As we can easily appreciate, IIT_w is IT, i.e., the thesis that in §3.2 I regarded as the primary target of Evans' *reductio*. But why should one consider IIT_w as the thesis that Evans aims to reject rather than IIT_r? I

contend that there are two basic reasons. The first reason consists in the fact that ITT_w seems to be a more suitable interpretation in light of Evans' own words. Indeed, keep in mind that Evans states that one of the two assumptions from which the thesis that he rejects arises is that 'the world might itself be vague', and not that our ways of referring to the world or our representations of it are vague. The second reason is that, if we interpret Evans' argument as aiming to disprove IIT_s rather than IIT_w, there is a simple way to block the argument that establishes conclusively that it poses no threat to the defenders of the view that identity is indeterminate in certain cases.

Let us discuss in further detail this second reason.[11] Suppose that the identity statement 'a=b' is indeterminate in truth-value only because 'a' is referentially indeterminate in virtue of semantic indecision: whereas 'b' determinately refers to the object b, there is no fact of the matter regarding whether 'a' refers to the object a or the object b. Now, even though 'a=b' is indeterminate in truth-value, one may appeal to the machinery of *precisifications* to circumvent this deficiency in the following manner: since a precisification consists in a way to interpret the language that eliminates the indeterminacy its expressions exhibit,[12] we may consider the interpretations of 'a=b' in which the statement is determinate. Given that 'b' refers determinately to b and 'a' could only refer to a or to b *ex hypothesi*, there are two admissible precisifications for the identity statement 'a=b' (call them P_1 and P_2 respectively). Granting that Ref('o') stands for the referent of 'o' and

[11] For a more detailed presentation of the argument given below, see Mena (MS), to whom I am greatly indebted. Cf. also Thomason (1982).

[12] For a more thorough discussion of the notion of precisification, which details its role in the framework of supervaluations, see Fine (1975).

that $[x\text{: } x{=}o]$ stands for the set of objects that are identical with o, we may describe these two precisifications as follows:

P_1: Ref('a')=a, Ref('b')=b, $[x\text{: } x{=}a]=\{a\}$, $[x\text{: } x{=}b]=\{b\}$

P_2: Ref('a')=b, Ref('b')=b, $[x\text{: } x{=}a]=\{b\}$, $[x\text{: } x{=}b]=\{b\}$

Consider now how Evans' argument fares within a model that includes only these two precisifications. Evans construes the initial supposition (1) as asserting that it is indeterminate whether '$a{=}b$' is true or not. Let us grant that, for any statement 'Φ' in the model, '$\bigtriangledown(\Phi)$' is true iff 'Φ' is true at some admissible precisification and false at another admissible precisification, and false otherwise. Thus, (1) is true iff '$a{=}b$' is true in at one precisification and false in another precisfication, and this is the case because '$a{=}b$' is true in P_2 and false in P_1.

Let us turn to (2). Given that (2) states that it is indeterminate whether b is identical to a or not, (2) is equivalent to the claim that it is indeterminate whether b belongs to the set of things that are identical with a or not. Let us grant that, for any object x in the model, it is indeterminate whether x belongs to the set of things that are identical to o iff there is an admissible precisification at which x belongs to the set of things identical to o and an admissible precisification at which x does not belong to the set of things identical to o. In light of this, (2) is true because b belongs to the set of things that are identical to a in P_2 but does not belong to the set of things that are identical to a in P_1.

When we consider the identity statement '$a{=}a$', we can appreciate without difficulty that it is true at both precisifications of the model. In light of

this, the statement '$\bigtriangledown(a{=}a)$' is false because it would only be true if '$a{=}a$' was true at some precification and false at another, and this is not the case. If we also grant that, for any statement Φ in the model, Φ is false iff $\neg\Phi$ is true, we can then conclude that '$\neg\bigtriangledown(a{=}a)$', which is (3), is true.

However, in the case of (4), a problem arises. Indeed, the lambda expression '$\lambda x[\bigtriangledown(x{=}a)]a$' is true iff a belongs to the set of things that are identical to a at one precisification and does not belong to the set of things that are identical to a at another precisfication, and this is the case because a belongs to the set of things identical to a at P_1 but does not belong to this set at P_2. And, since '$\lambda x[\bigtriangledown(x{=}a)]a$' is true, its negation '$\neg\lambda x[\bigtriangledown(x{=}a)]a$', which is (4), turns out to be false. Consequently, a rather simple model generated by the hypothesis that 'b' refers determinately to b and that it is indeterminate whether 'a' refers to a or b invalidates Evans' argument.

Several authors (e.g., Thomason 1982) have remarked that the reason that accounts for the invalidity of the inference from (3) to (4) in this model is the occurrence of a fallacy similar to the one displayed by the following inference:

(CJ) It is contingent that the number of Justices is odd.

$\therefore$(JC) The number of Justices is such that it is contingently odd.

This inference is fallacious because 'The number of Justices' is a non-rigid designator that cannot be freely moved past the contingency operator. Indeed, 'The number of Justices' makes CJ true because, when it occurs within the scope of the contingency operator, it refers to distinct numbers in different counterfactual situations but, when it occurs outside of the scope of

the contingency operator, it refers to the number nine, which is necessarily odd —thus making JC false. Similarly, the problem with the inference from (3) to (4) under the assumption that 'a' is referentially indeterminate is that, in spite of its *referential flexibility*, any interpretation of 'a' makes true '$\neg\bigtriangledown(a{=}a)$' because both occurrences of 'a' fall under the scope of the indeterminacy operator but that, since the occurrence of 'a' outside of the scope of '$\bigtriangledown$' in (4) determinately refers to a whereas the occurrence of 'a' governed by '$\bigtriangledown$' admits distinct interpretations, '$\neg\lambda x[\bigtriangledown(x{=}a)]a$' turns out to be false.

Evans' argument may then be blocked effectively if it is interpreted as targeting IIT_s. However, such an interpretation is problematic because it suggests that Evans completely overlooked the fallacy. A far more charitable (and plausible) interpretation of Evans' argument consists in maintaining that Evans grants that there are indeterminate identity statements that are so in virtue of semantic indecision and that what he really intends to show is that the indeterminacy in truth-value of identity statements can only be due to referential indeterminacy that stems from semantic indecision and not from any form of worldly indeterminacy —i.e., that he intends to disprove IIT_w. Following Lewis (1988), who claims that this is the interpretation originally intended on the grounds of an exchange of letters with Evans himself, I then construe Evans' argument as a *reductio* of IIT_w.

Construing Evans' argument as a *reductio* of IIT_w has the following consequence: given that Evans aims to show that the view that identity statements can be indeterminate in truth-value exclusively in virtue of worldly indeterminacy is incoherent, a fair-minded interpretation of the argument requires

eliminating the occurrence of referential indeterminacy due to semantic indecision in the supposition for *reductio*. This result is usually guaranteed by assuming that 'a' and 'b' are both *precise designators*.[13] Indeed, this assumption allows us to ensure that, if one attempts to block Evans' argument by diagnosing the occurrence of some form of referential indeterminacy either in the initial supposition or any of the subsequent steps, this referential indeterminacy must not be an accidental feature of our language that occurs against the background of a determinate world but rather the upshot of some form of worldly indeterminacy.[14] Accordingly, I take for granted in the following sections, as all of the authors whose views I discuss below do, that the singular terms 'a' and 'b' that Evans employs to run his argument are precise designators that denote determinately one object.

4.4 Two objections from Parsons and Woodruff

After presenting two interpretations of Evans' argument in the previous section and arguing that Evans really intended to show that identity statements cannot be indeterminate in truth-value in virtue of worldly indeterminacy, I can now turn to rehearse and critically assess several objections

[13]The notion of a precise designator closely mirrors that of a rigid designator in the following way: just as a rigid designator is a referring expression that denotes the same object at all the counterfactual situations where that object exists and does not denote anything else at counterfactual situations where that object does not exist, a precise designator is a referring expression that denotes the same object at all the precisifications where that object exists (i.e., all the admissible precisifications) and does not denote anything else at the precisifications where the object does not exist (i.e., the inadmissible precisifications). For a slightly different but equivalent characterization of the notion, see Garrett (1991: 342).

[14]This is precisely the argumentative line that Williams (2008) exploits. For further details on this argumentive line and some critical discussion of it, see §4.7 below.

that have been addressed to that interpretation of the argument. In this section, I specifically consider two objections that Parsons and Woodruff (1995) have laid down —one of which stems from Evans' use of lambda abstracts to express properties in the argument whereas the other concerns the use of contraposition. Here is in further detail how I proceed below. In §4.4.1, I outline briefly certain background assumptions that Parsons and Woodruff endorse regarding the nature of wordly indeterminacy, the expression of indeterminacy in logic and the characterization of identity. Then, I introduce in §4.4.2 both objections to the argument. Finally, in §4.4.3, I show how both objections can be handled successfully by a defender of Evans' argument.

4.4.1 Some background assumptions

After reminding us that there are several puzzles and thought experiments that provide some initial motivation for the thesis that identity is indeterminate in certain cases in virtue of a worldly source, Parsons and Woodruff suggest that the worldly source in question consists, not in vague objects, bur rather in vague *states of affairs*. Given that states of affairs are characterized in terms of the possessings of certain properties by certain objects and the standing-ins of certain relations between certain objects, the worldly indeterminacy they consider consists in whether certain properties and/or relations are instantiated or not. The advantage of such a view is that it enables them to subsume the wordly indeterminacy due to vague objects under a broader framework: if an object is vague because it has fuzzy boundaries, this may be also characterized in terms of the absence of a matter of fact about its location. The view also matches well the characterization of identity that

they endorse, and that I describe concisely below.

In addition to the characterization of worldly indeterminacy in terms of the notion of vague states of affairs, Parsons and Woodruff also defend a view that precludes higher-orders of indeterminacy. Indeed, given that they (1995: 173) maintain that, for any object x and any property P, x may have P or lack P or the world may not determine either one of these and that 'exactly one of these must hold', it is patent that their framework does not allow higher-orders of indeterminacy. In light of this, Parsons and Woodruff presuppose a classical conception of indeterminacy in which the existence of borderline cases is acknowledged, but these are considered to make up a definite class with precise boundaries.[15]

A third assumption that Parsons and Woodruff adopt concerns the characterization of identity. For them (*ibid.*), the best way to characterize identity of, say, a and b, is not *via* the interchangeability between the names respectively used to refer to a and b (e.g., 'a' and 'b') but rather *via* the 'indistinguishability of a and b in terms of the actual properties and relations they have or stand in.' The rationale for this is twofold: on one side, there are terms that do refer to the same object but are not interchangeable in virtue of the existence of opaque contexts and, on the other side, there are languages in which the names of distinct objects may be intersubstitutable because the languages in question lack enough expressive ressources to effectively distinguish the objects. As I mentioned earlier, this characterization of identity matches well the conception of worldly vagueness that Parsons and Woodruff defend and, by combining them, it gives rise to a test for

[15]For a criticism of the classical conception of indeterminacy, see Sainsbury (1991). Further discussion of this issue can also be found in Wright (1992) and Hyde (1994).

indeterminate identity that is described as follows:

> That test is whether a and b share all the same properties and stand in all the same relations. Whether this holds in any particular case may be undetermined by the facts. If so, a and b will actually be neither determinately identical nor determinately distinct. This will be a case of indeterminacy of identity. The framework of properties and relations sketched above is neutral about whether this ever actually happens. (1995: 174)

Having introduced the test for indeterminate identity, Parsons and Woodruff make a further assumption about what asserting a sentence amounts to. Indeed, since Evans aims to give a *reductio* of IIT_w, he is required to assert, for the purposes of the *reductio*, the thesis from which he intends to derive a contradiction and, in virtue of this, an account of what the assertion a sentence consists in is needed. The suggestion that Parsons and Woodruff present to satisfy this demand goes as follows: to assert a sentence S boils down to present it as true. In light of this, if one asserts '$\bigtriangledown(a{=}b)$' and the sentence turns out to be false, then it was mistake to assert it in the first place. This assumption is very important because, without it, the formal derivation of '$\neg(a{=}b)$' from '$\bigtriangledown(a{=}b)$' remains a mere syntactic exercise without philosophical import.

Let us recap now. Prior to assessing Evans' argument, Parsons and Woodruff introduce four basic assumptions: (1) an interpretation of worldly indeterminacy in terms of a framework of vague states of affairs, which are constituted by objects, properties and relations, (2) a classical conception

of indeterminacy that precludes the existence of higher-order indeterminacy, (3) a characterization of identity in terms of indistinguishability of properties and relations and (4) the thesis that asserting a sentence S requires presenting it as true. With these assumptions spelled out, we may turn now to consider the objections.

4.4.2 The two objections

Though some commentators (e.g., Williams 2008: 137fn8) occasionally refer to the defense of IIT_w mounted by Parsons and Woodruff as being grounded on just one objection against Evans' argument, it is important to observe that Parsons and Woodruff really present *two* objections, which I rehearse in detail further down. For the sake of expediency, I just label them as 'the ontological objection' and 'the logical objection' respectively.

4.4.2.1 The ontological objection to Evans' argument

Parsons and Woodruff observe that, if we keep in mind both the formal apparatus that Evans deploys in his argument and the interpretation that he originally intended for it, it becomes clear that there are a series of constraints that the argument should abide by to be valid. One constraint arises in connection with the use of lambda expressions in (2) and (4). Indeed, it is well known that, though our language can be enriched by adding to it predicates generated through the application of lambda abstraction on every single name it includes, there are serious restrictions on the interpretations that some of these lambda predicates admit. In particular, the existence of Russell's paradox shows that not every predicate generated by lambda

abstraction can be considered to express a property, lest we want to admit the existence of entities such as the property of not having any property at all.

Besides this constraint, Parsons and Woodruff remark that there is another constraint that Evans' argument should respect in virtue of its intended interpretation. Here is how they formulate it:

> Suppose our language contains an indeterminacy operator '$\bigtriangledown$', which obeys these laws:
>
> $\bigtriangledown(S)$ is true iff S is lacking in truth-value.
>
> $\bigtriangledown(S)$ is false iff S is true or is false.
>
> Then one *cannot* take for granted that lambda abstracts that bind variables in contexts governed by '$\bigtriangledown$' stand for properties *and also* fully satisfy lambda abstraction. (1995: 175)

Since the formulation of the constraint is rather condensed, let me unpack it by considering the case that Evans uses to run his argument. Suppose that the identity statement 'a=b' is indeterminate in truth-value. This entails, in virtue of the rules that govern the indeterminacy operator '$\bigtriangledown$', that '$\bigtriangledown(a$=$b)$' is true. Now, if lambda abstraction could be used unrestrictedly on '$\bigtriangledown(a$=$b)$' to generate a lambda abstract that expresses a property, the property in question would be the property *being indeterminately identical with* a. But, given that '$\bigtriangledown(a$=$b)$' expresses the absence of a matter of fact regarding whether b is identical with a, there is a *prima facie* doubt that arises concerning whether this absence of a matter of fact determines a fur-

ther fact that involves the property *being indeterminately identical to a*.[16] Consequently, since there are reasons to doubt that the lambda abstract in (2) expresses a genuine property while satisfying fully the principle of lambda abstraction, Parsons and Woodruff maintain that Evans cannot assume that this is the case without begging the question.

Parsons and Woodruff further elaborate on this issue by arguing that, since Evans does not have an argument that shows that the lambda abstracts that he employs in his *reductio* both fully satisfy lambda abstraction and stand for genuine properties, he has two alternatives to avoid making the question-begging assumption. Unfortunately, none of the alternatives is satisfactory because, as they argue in the following passage:

> One is to assume that the principle of abstraction always holds for predicates, but to leave it open whether the resulting predicates stand for properties. If we do this, the argument begs the question at step (5). The other option is to take for granted that abstracts stand for properties, but reject the principle of abstraction as always providing conditions under which such properties hold of objects. Then the abstraction steps (2) and (4) may be unjustified. (1995: 177)

It important to highlight that the focus of this objection is not a semantic issue, but rather an ontological one. Indeed, Parsons and Woodruff do not challenge the use of lambda abstraction by arguing that the lambda predicates are meaningless. They admit that the lambda predicates are significant

[16] In particular, Keefe (1995: 188) argues that, whereas some statements describe a mode in which an object has a property, '$\triangledown(a{=}b)$' expresses simply that there is no fact of the matter whether b has the property in question.

expressions, but they question whether these lambda predicates refer to genuine properties. Indeed, since the argument appeals to the contrapositive of Leibniz's Law to derive (5) from (2) and (4) and Leibniz's Law states a condition on the *properties* that any objects a and b have if they are identical (and not merely on the predicates that are satisfied by a and b or on the names that are used to refer to a and b respectively), the inference from (2) and (4) to (5) presumably requires to be valid that the lambda predicate '$\lambda x[\bigtriangledown(x{=}a)]$' refers to a property. Consequently, since Evans cannot justify his assumption that the lambda predicate '$\lambda x[\bigtriangledown(x{=}a)]$' that occurs in (2) and (4) fully satisfies lambda abstraction and expresses a genuine property without begging some question of ontology, Parsons and Woodruff conclude that the *reductio* fails.

4.4.2.2 The logical objection to Evans' argument

After arguing that Evans begs a question when he assumes that the lambda predicate '$\lambda x[\bigtriangledown(x{=}a)]$' in his argument fully satisfies lambda abstraction and expresses a real property, Parsons and Woodruff raise a different objection. They point out that, though Evans seems to treat in the argument the lambda predicate '$\lambda x[\bigtriangledown(x{=}a)]$' as expressing a property, Evans could argue (in a Quinean fashion) that this does not commit him to the existence of the property *being indeterminately identical to a* because the argument does not involve any quantification over properties. Indeed, even though Evans could use (2) and (4) to derive $(\exists\Phi)[\Phi(b)\ \&\ \neg\Phi(a)]$ by conjoining first (2) and (4) and then performing existential generalization on the occurrences of the predicate '$\lambda x[\bigtriangledown(x{=}a)]$' in the resulting statement, Parsons and Woodruff ob-

serve that he moves directly from (2) and (4) to (5) just by invoking Leibniz's Law without quantifying over properties.

This is precisely where Parsons and Woodruff introduce a second, two-step objection to Evans' argument. The objection involves two steps because Parsons and Woodruff first stress that, although Evans invokes Leibniz's Law to obtain (5), he employs its contrapositive. Having done that, they argue that the use of Leibniz's Law contrapositive in Evans' argument is illegitimate because the contrapositives of certain arguments and principles that are truth-preserving in classical logic turn out to be invalid in contexts that sanction the possibility of indeterminacy in truth-value:

> In classical logic without truth-value gaps a principle and its contrapositive are equivalent. But if there is a possibility of truth-value gaps, a principle and its contrapositive are not necessarily equivalent, and presuming that they are equivalent begs a question of logic that arises independently of any question about identity. (1995: 177-178)

To illustrate this, Parsons and Woodruff ask us to consider the following argument schema:

$$\text{TNI}\frac{A}{\therefore \neg \bigtriangledown A}$$

This argument schema clearly seems to be truth-preserving, but that is not the case for the contraposed argument schema:

$$\text{INT}\frac{\bigtriangledown A}{\therefore \neg A}$$

Just as we did previously in the case of Evans' assumption that the lambda predicate expresses a property, it is important to notice that the focus of the current objection is not an issue of semantics. In fact, it is also not a issue of ontology but an issue of logic. Let me spell out this in further detail. According to Parsons and Woodruff, Evans invokes a well-known principle of logic (i.e., Leibniz's Law) to justify the derivation of (5) from (2) and (4), but he really uses its contrapositive. However, he is not entitled to do this because, although the validity of a principle or an argument entails the validity of its contrapositive in classical logic, his argument requires being interpreted in a non-classical logic.

Parsons and Woodruff justify this contention by calling attention to the fact that Evans supposes, for the purposes of carrying out his *reductio*, that some identity statements are indeterminate in truth-value. Thus, considering that Evans is forced to adopt, in virtue of this supposition, a logic in which 'lack of truth-value is taken seriously' (1995: 174), some of the common principles of classical logic may break down in this logic. This is specifically what occurs in the case of the principle stating that the validity of a certain principle or argument entails the validity of its contrapositive, as the fact that the argument TNI is valid but not its contrapositive INT shows. To conclude, Parsons and Woodruff argue that, since Evans employs the contrapositive of the Indiscernibility of the Identical in his *reductio* because he assumes that it is as valid as Indiscernibility of the Identical is but that he is not entitled to make that assumption because the contrapositives of certain valid principles turn out to be invalid in a non-classical logic that sanctions indeterminacy of truth-value, he also begs a question of logic in addition to begging a question

of ontology.[17]

4.4.3 Responses to the objections

4.4.3.1 Response to the ontological objection: the nature and role of properties

As we have previously seen in §4.4.2.1, the gist of the ontological objection is that Evans assumes without argument that the lambda predicate '$\lambda x[\bigtriangledown(x{=}a)]$' both satisfies fully lambda abstraction and refers to a property. Parsons and Woodruff maintain that such an argument is required because there is a *prima facie* doubt regarding whether the predicate can fulfill both demands motivated by the following difficulty: if we grant that the predicate '$\lambda x[\bigtriangledown(x{=}a)]$' expresses a property, it would be the property *being indeterminately identical with a* but, since the predicate is also generated by lambda abstraction from (1) —which expresses the absence of a matter of fact regarding whether a is identical with b, we are seemingly forced to accept that the absence of a matter of fact gives rise to a further fact which involves the abovementioned property —and this is rather counterintuitive.

To answer this objection, let us recall the first assumption that Parsons and Woodruff make: they construe worldly indeterminacy in terms of a metaphysical framework of properties and relations. To be more accurate, they assume that worldly indeterminacy consists in vague states of affairs and that states of affairs are nothing but the havings of certains properties by

[17]Even though Parsons and Woodruff raise this objection in connection with Evans' use of the contrapositive of the Indiscernibility of the Identical to obtain (5) from (2) and (4), they could also in principle raise it in connection with his use of the contrapositive of property concretion to derive (4) from (3).

certain objects and the standing-ins of certain objects in certain relations. In light of this, it is patent that Parsons and Woodruff embrace implicitly a certain conception of the nature of properties and relations: they are the *constituents* of states of affairs that are not objects themselves.

I maintain that it is because Parsons and Woodruff embrace this conception of properties that they are led to believe that Evans needs to give an argument to justify the assumption he makes about the lambda predicate. Indeed, given this particular conception of properties, it is virtually impossible to explain how a vague state of affairs (e.g., the absence of a matter of fact regarding whether an object o has a certain property F) can determine a further state of affairs (e.g., o's having the property *being indeterminately* F) if we subscribe, as I do, to Russell's thesis (1923: 85) that 'things are what they are, and there is an end to it.' Indeed, if there is no fact of the matter regarding whether o is F and that is just the way things are, arguing that there is a further state of affairs —o's being indeterminately F— that is determined by the absence of the matter of fact seems to be just a way of *misdescribing* the way things are.

Now, the conception of properties as constituents of states of affairs corresponds to a view according to which the defining function or role of properties is to provide a reductive account of states of affairs. But this role is not the only one that is usually attributed to properties. In fact, properties are used for such a wide variety of purposes in philosophy that I cannot but agree with the following observation made by Lewis:

> (...) we have the word 'property', introduced by way of a varied repertory of ordinary and philosophical uses. The word has

> thereby become associated with a role in our commonsensical thought and in a veriety of philosophical theories. To deserve the name of 'property' is to be suited to play the right theoretical role; or better, to be one of a class of entities which together are suited to play the right role collectively. But it is wrong to speak of *the* role associated with the word 'property', as if it were fully and uncontroversially settled. The conception is in considerable disarray. (1986: 55)

Being the constituents of states of affairs is one of many *metaphysical* roles that is traditionally attributed to properties. However, they are also attributed a *semantic* role by many philosophers —in particular, they are identified with the meanings of predicates such as '(...) is tall' and of abstract singular terms such as 'whiteness'. Given that Lewis persuasively argues in the abovementioned passage that there is no definitive or ultimate role that singles out properties once and for all but rather a variety of different roles, I maintain that, although Evans' *reductio* is indeed invalid if we assume a metaphysical conception of properties such as the one Parsons and Woodruff endorse, this does not undermine the *reductio* because nothing forces Evans to adopt this particular conception. All that Evans needs to run the argument is a conception of properties robust enough to allow him to claim that (2) and (4) respectively follow from (1) and (3) and that they conjointly yield (5) in virtue of the contrapositive of Leibniz's Law —and the best candidate for this is a semantic conception of properties that identifies them with the meanings of predicates and abstract singular terms. In fact, since Parsons and Woodruff admit that the lambda predicate '$\lambda x[\bigtriangledown(x{=}a)]$'

is not meaningless, they too have the basic tools to accomodate an ontologically lean interpretation of Evans' argument (which is presumably the one that Evans himself intended).[18] Thus, the ontological objection to Evans' argument is unsuccesful because it ultimately targets a strawman version of the argument.

Before considering the logical objection, I want to offer a brief and tentative diagnosis of why Parsons and Woodruff were led to criticize a strawman version of the argument. Recall that Evans claims that one of the views that give rise to the thesis he wants to reject is the view that 'the world itself might be vague' —i.e., that things themselves might be indeterminate rather than the words we use to refer to them or our representations of them. In light of this, since one rather plausible hypothesis concerning the basic structure of the world is that it is made of states of affairs and that these states of affairs are constituted by objects, properties and relations, it also seems plausible to interpret Evans as embracing an ontology of properties-*qua*-constituents-of-states-of-affairs to flesh out the view that the world might be vague in terms of the possibillity that some of these properties are indeterminately instantiated.

But assuming that there are vague states of affairs is not the only way to interpret the view that the world might be vague. To appreciate this, suppose that we share Quine's (1969: 19-23) strong distaste for properties, propositions and other abstract entities that lack identity conditions, and are

[18]Indeed, there is a sense in which the semantic conception of properties does not call for any ontology in particular because, even though some semantic theories identify meanings with platonic universals or with sets of entities, others, as Oliver (1996: 16) remarks, leave open the question about their nature and merely affirm that a property is whatever plays a certain semantic role.

only willing to countenance concrete particular objects and classes made out of those in our ontology. Even in this austere framework, we can still make sense of the view that the world might be vague in terms of *indeterminacy of class membership*: some objects are such that it is indeterminate whether they belong to a certain class or not. Thus, even if the metaphysical framework of properties and relations that Parsons and Woodruff assume provides a natural way to understand the view that the world might be vague, Evans is not compelled to accept it since there are many other plausible ways to flesh out the view that reject the framework as a whole.

4.4.3.2 Response to the logical objection

The main thrust of the logical objection consists in arguing that, since Evans employs the contrapositive of the Indiscernibility of the Identical (rather than the Indiscernibility of the Identical itself) in his *reductio* and the contrapositives of several truth-preserving arguments and principles are invalid in contexts that sanction lack of truth-value, Evans begs a question of logic when he assumes without argument that he can use the contrapositive of Leibniz's Law to derive (5) from (2) and (4). To respond to this objection, I first want to examine the assertion made by Parsons and Woodruff according to which the logic that Evans has to assume to carry out the *reductio* is one in which 'lack of truth-value is taken seriously' and, having done that, I take a close look at the case that Parsons and Woodruff introduce to illustrate the failure of the principle that the validity of an argument entails the validity of its contrapositive in a non-classical logic.

I have argued earlier in §4.2 that, given the paucity of the assumptions and

the formal machinery used in Evans' argument, nothing forces him to adopt a special 'logic of indeterminacy' to run the *reductio*. This assertion clashes, of course, with the claim that Evans' argument calls for a non-classical logic that takes lack of truth-value seriously. But what motivations are there to maintain that Evans' argument calls for this sort of logic? Even though it is patent that Evans introduces a sentential operator '$\bigtriangledown$' to express indeterminacy in truth-value, it is crucial to bear in mind that the introduction of the operator is not accompanied in Evans' paper by any semantics or specific rules of inference for sentences involving the operator. This has been considered implicitly as an oversight by several authors who have, like Parsons and Woodruff, stepped in and provided semantics to interpret sentences that involve '$\bigtriangledown$'.[19] However, I contend that this lack of semantics or specific rules of inference governing '$\bigtriangledown$' in Evans' paper is deliberate, and that it is due to the fact that Evans does *not* take seriously the thesis that identity statements may lack truth-value in virtue of some form of worldly indeterminacy. He introduces the operator '$\bigtriangledown$' as part of a *linguistic framework* the purpose of which is to enable him to talk about the indeterminacy of truth-value of some identity statements that is due to worldly indeterminacy, but without being committed to it.[20]

In addition to the lack of semantics or specific rules of inference for sentences governed by '$\bigtriangledown$', another indication of the fact that Evans does not

[19] In addition to Parsons and Woodruff, Van Inwagen (1990) introduces some rather elaborate semantics for sentences involving the indeterminacy operator. An extended discussion of Van Inwagen's treatment of Evans' argument can be found further down in §3.5

[20] The classical presentation and defense of linguistic frameworks as tools that enable us to accept certain systems of entities without being committed to them can be found in Carnap (1950).

take seriously IIT_w comes to light when we consider the fourth assumption that Parsons and Woodruff introduce to understand Evans' argument — namely, the thesis that asserting a sentence S consists in presenting it as true. If Evans asserts (1) to lay down the first step of the *reductio*, he then presents (1) as true in virtue of the assumption. But isn't this problematic in the sense that to present something as true is similar, as Dummett (1958: 142) suggests, to present something as a winning move when playing a game? Indeed, if this is the case then, just as I am required to uphold a winning move in a game of chess from a challenge, I am also required in principle to uphold whatever I present as true from a challenge. However, Evans' attitude does not seem to fit this model because, although he presents (1) as true, he does this with the explicit purpose of showing that the thesis is incorrect.

The solution to this apparent tension consists in acknoweldging that, though presenting a sentence as true often involves being disposed to uphold it, that is not always the case. There are certain circumstances in which, even if we present a sentence or a thesis as true, we are not really disposed to uphold it. These circumstances typically involve according to Yablo (1998: 234) "make-believe games [which] are the paradigm activities in which we 'assent' to sentences with little or no regard for their actual truth-values." Now, considering that the introduction of Carnap's linguistic frameworks is quite similar to playing a make-believe game in the sense that none of these two activities require endorsing a belief that things are such-and-such[21] and that the best way to interpret Evans' attitude vis-à-vis (1) is to maintain that he plays a make-believe game, it is rather plausible to maintain that

[21] In fact, Yablo argues in the aforementioned paper that Carnap's linguistic frameworks are a kind of make-believe games.

Evans does not take seriously IIT_w and that he introduces the sentential indeterminacy operator '$\bigtriangledown$' as part of a linguistic framework.

If what I argued in the preceding paragraphs is accurate, Evans' argument does not call for a special 'logic of indeterminacy' in spite of the fact that it involves an operator '$\bigtriangledown$' that expresses indeterminacy in truth-value. This result is very important because it enables us to block not only the logical objection raised by Parsons and Woodruff, but a family of objections to Evans' argument that have the same underlying structure.[22] The first stage of all these objections consists in claiming that, to address properly issues of worldly indeterminacy, a special non-classical logic is required. There are divergences concerning which logic is needed, of course: whereas some suggest it must be a logic that allows truth-value gaps (e.g., a supervaluationist logic), others advocate a logic that allows truth-values in addition to truth and falsity (e.g., a trivalent or fuzzy logic). The next stage consists in arguing that, since Evans intends to show that identity statements cannot be indeterminate in truth-value in virtue of worldly indeterminacy, his argument has to be assessed in whatever non-classical logic allows us to make best sense of indeterminacy of truth-value due to worldly indeterminacy. Finally, it is shown that Evans' argument founders in the non-classical logic in question because at least one of the inferences turns out to be invalid.

Now, granting that Evans's argument does not call for a 'logic of indeterminacy' insofar as the use of '$\bigtriangledown$' in the *reductio* is just a manner of expressing the indeterminacy in truth-value of certain identity statements due to worldly indeterminacy in the context of a Carnapian linguistic framework,

[22] In addition to Parsons and Woodruff, Broome (1984), Johnsen (1989) and Tye (1990) address the same type of objection to Evans' argument.

the primary motivation to develop non-classical models for the argument is greatly abated because assessing the argument in these models would clash with Evans' intentions. Indeed, if Evans introduces '$\bigtriangledown$' as a mere *linguistic form* that is used to talk about indeterminacy in truth-value due to worldly indeterminacy without being committed to it, the presence of '$\bigtriangledown$' does not force him, *pace* Parsons and Woodruff, to take indeterminacy of truth-value seriously. Consequently, using a non-classical model that sanctions either truth-value gaps or additional truth-values to interpret the *reductio* contradicts Evans' intentions because the use of any of these non-classical models amounts to take indeterminacy of truth-value due to worldly indeterminacy seriously.

Another parallel issue that also casts serious doubts on the logical objection is the following. Just as Evans' intentions involve using '$\bigtriangledown$' to talk about indeterminacy in truth-value due to worldly indeterminacy in a non-committal way, he presumably intends to use '=' in his argument to refer to identity given that he aims to show that the thesis that some *identity* statements are indeterminate in truth-value in virtue of worldly indeterminacy is false. However, if this is the case, allowing the argument to be interpreted in a non-classical model (i.e., a model that sanctions identity statements that either lack truth-value or have a truth-value that is neither truth nor falsity) on the grounds that the argument presupposes a logic that takes seriously indeterminacy of truth-value due to worldly indeterminacy contradicts Evans' intention of using '=' to denote identity. Williamson has incisively argued for this claim in the following passage:

> If '=' were interpreted faithfully, the truth-value of 'a=b' would

> depend simply on the identity or otherwise of the denotation of a in the model with the denotation of b in the model: the formula would be true in the model if they were identical and false in the model if they were not identical. But that is not what we find. Rather, the model theory allows us to assign straightforwardly distinct objects as the denotations of a and b respectively in a model, but still to treat the formula 'a=b' as half-true or at least not false in that model. (2002: 282)

Given that interpreting Evans' argument in a non-classical model that sanctions identity statements with a truth-value gap or that admit additional truth-values entails contradicting Evans' intentions vis-à-vis the interpretation of '=', it seems that interpreting the argument in such a model boils down to change the subject. Indeed, if we allow Evans' argument to be interpreted in a non-clasical model in which identity statements of the form 'a=b' may be half-true or not false in spite of the fact that 'a' and 'b' are assigned different denotations, it is clear that Evans' argument may turn out to be invalid in such a model, but that constitutes no objection against it because the model is unfaithful to the intended meaning of '='. The model presumably captures some other relation that is similar to identity in many respects. However, this is not good enough to undermine the *reductio*: an advocate of Evans may well respond to those who attempt block the argument by showing that it is invalid in some non-classical model that they are compelled to change the subject because the strictures that their non-classical model imposes on the meaning of '=' are such that '=' does not longer refer

to identity, as Evans intended.[23]

I have argued in the previous paragraphs that Evans is not forced to assume a non-classical logic to carry out his argument against indeterminate identity because the only element that could justify the use of a non-classical model to interpret the argument —namely, the sentential indeterminacy operator '$\bigtriangledown$'— is introduced in the context of a Carnapian linguistic framework as a linguistic form used to talk about indeterminacy of truth-value due to worldly indeterminacy in a noncommittal way. Moreover, I have also argued that there is an further reason to maintain that Evans' argument does not call for a non-classical model: these models are ultimately unfaithful to the meaning of '=' that Evans intended in his argument. I want to consider now in more detail the example that Parsons and Woodruff use to illustrate the logical objection.

Parsons and Woodruff maintain that, whereas the argument TNI is truth-preserving, the contraposed argument INT is not. While I am in total agreement with their assessment of INT, I have some serious reservations vis-à-vis TNI, which cast doubt on the example they use to illustrate the logical objection. Let me explain. TNI asserts that $\neg\bigtriangledown A$ can be inferred from A. Since Parsons and Woodruff do not mention any restrictions to the application of TNI, they presumably hold that it is valid for any statement A whatsoever. But suppose A abbreviates $\bigtriangledown B$. If that is the case and TNI is indeed valid for any A, then Parsons and Woodruff have to accept that $\neg\bigtriangledown\bigtriangledown B$ can be derived from $\bigtriangledown B$. This is, of course, consistent with the second assumption they make —namely, the adoption of a classical conception of indeterminacy

[23]For a classical illustration and defense of the thesis that a change of logic entails a change of subject, see Quine (1986).

that precludes the existence of higher-orders of indeterminacy. But how plausible is this assumption in light of the fact that Parsons and Woodruff also assume a metaphysical framework of states of affairs to account for worldly indeterminacy?

I contend that the two assumptions are at odds with each other because the metaphysical framework of states of affairs that Parsons and Woodruff introduce to account for worldly indeterminacy calls for higher-orders of indeterminacy. To appreciate this, consider the following argument. Suppose that the world is such that it does not determine whether an object o (e.g., an electron) instantiates a property P (e.g., *being at a certain location l*). Let us also grant that we refer to this lack of a fact of the matter as $\bigtriangledown Po$. This lack of a fact of the matter may be determinate or indeterminate. If it is determinate (as Parsons and Woodruff presumably maintain), then it should in principle be determined by the world. However, since the world does not determine whether o instantiates P or not *ex hypothesi*, it is impossible to explain how the lack of a fact of the matter regarding whether o instantiates P could account for its own determinacy. Accordingly, the lack of a fact of the matter regarding whether o instantiates P is indeterminate. But this boils down to affirm $\bigtriangledown\bigtriangledown Po$, and thus to acknowledge the existence of higher-orders of indeterminacy.[24]

In light of the previous argument, I believe that we have very strong motivations to question the assumption that Parsons and Woodruff make concerning the adoption of a classical conception of indeterminacy. However,

[24]Several authors consider the existence of higher-orders of indeterminacy to be a distinctive feature of indeterminacy itself. For illustrations of this view, see Fine (1975: 287) and Akiba (2002: 83).

if we put the assumption in question and thus acknowledge higher-orders of indeterminacy, the thesis that the argument schema TNI is truth-preserving is severely compromised. This undercuts decisively the effectiveness of the example that Parsons and Woodruff use to illustrate the logical objection they address to Evans' argument. Consequently, given that the assertion that Evans is forced to assume a non-classical logic to establish his *reductio* may be succesfully resisted and that the case that Parsons and Woodruff use to illustrate their objection is quite problematic, I conclude that the logical objection does not put at risk Evans' argument.

4.5 Van Inwagen's objection

In the present section, I consider thoroughly and respond to Van Inwagen's objection. To do this, I proceed in the following way. In §4.5.1, I rehearse briefly the main elements of the background discussion that motivates Van Inwagen's response to Evans. In particular, I introduce the metaphysical question about composition that Van Inwagen (1990) addresses and I show why the solution that he defends vis-à-vis this question leads him to acknowledge that, in some cases, certain identity statements are indeterminate in truth-value in virtue of worldly indeterminacy —an admission that puts him at odds with Evans. After doing this, I present in §4.5.2 the objection that Van Inwagen raises against Evans' *reductio*, which consists in developing a semantics that includes all the syntactic components used by Evans and showing that the *reductio* turns out to be invalid in this interpretation. Finally, I draw attention in §4.5.3 to a number of difficulties that

cast serious doubts on Van Inwagen's proposal.

4.5.1 The background discussion: the special composition question

As I previously mentioned, Van Inwagen's objection to the *reductio* arises in the context of a discussion concerning the composition question. Van Inwagen stresses the importance of addressing this question by observing that, although some philosophers admit the existence of composite material objects (i.e., material objects which have as their proper parts other objects that compose them), they are sometimes silent concerning how these material objects can be composed of other material objects.[25] But this silence is, at least initially, rather unsatisfactory because, without an explanation of why composition occurs in the case of composite material objects, we face some problematic puzzles such as the following (which assumes that a cat is a composite material object):

> Consider a cat that is composed at t of certain atoms arranged in a certain way; it is at least logically possible for those very same atoms to be arranged in the same way at some later time and then to compose a different cat; but that apparently implies that certain things can compose a material object at t, and, even though arranged in precisely the same way, compose a distinct material object later. How could that be? (1990: 18)

[25] In some cases, this is because they maintain that there is no explanation of why certain material objects compose a further material object: it is just a brute, primitive fact that they do compose it. For a defense of this view, see Markosian (1998).

In light of this, we require to have a view that explains under what circumstances composition takes place. This view would constitute an answer to what Van Inwagen calls the 'Special Composition Question' (SCQ). Considering that the main goal of Van Inwagen in *Material Beings* is to develop and argue for a particular response to the SCQ, let me briefly rehearse a standard formulation of the SCQ before presenting and discussing Van Inwagen's response to it as well as some of its implications. In order to formulate the SCQ, Van Inwagen first introduces plural variables such as '$\mathcal{X}$s', '$\mathcal{Y}$s'... to refer collectively to pluralities of entities. Subsequently, he characterizes the notion of *overlap* in terms of the primitive mereological notion of *part* in the following way:

> Overlap: $(\forall x)(\forall y)[x$ overlaps y iff $(\exists z)(z$ is a part of x and z is a part of $y)]$

With the help of plural variables and the notions of part and overlap, Van Inwagen (1990: 29) then defines the dyadic predicate 'compose' in the following terms:

> Composition: The $\mathcal{X}$s compose y iff (i) the $\mathcal{X}$s are all parts of y and (ii) no two of the $\mathcal{X}$s overlap with each other and (iii) every part of y overlaps at least one of the $\mathcal{X}$s.

Once the notion of composition is in place, Van Inwagen has all the elements needed to formulate the SCQ:

> SCQ: What are the circumstances (if there are any) in which certain entities (i.e., the $\mathcal{X}$s) compose a further entity y?

After formulating the question, Van Inwagen examines several possible responses that can be given to it. In particular, Van Inwagen considers a series of moderate views, which involve maintaining that composition requires that some kind of contact or physical bonding takes place between the $\mathcal{X}$s, as well some extreme views, which involve claiming either that there are no circumstances in which any $\mathcal{X}$s compose a further entity y (nihilism) or that any $\mathcal{X}$s whatsoever compose in every circumstance a further entity y (universalism). Since all these responses to the SCQ are deemed to be problematic for a number of reasons, Van Inwagen (1990: 82) then suggests a further view, which is his proposed answer the SCQ:

> Life: $(\exists y)$(the $\mathcal{X}$s compose y) iff the activity of the $\mathcal{X}$s constitutes a life.

Let me flesh out Van Inwagen's proposal in further detail by explaining what he means by 'activity constituting an event' and 'life'. The notion of an activity that constitutes an event is basically accounted for through a series of examples: for instance, if my signing a check and handing it to my landlord can be considered as an activity, this activity constitutes paying my rent and if the reaction of dioxygen molecules with methane can be considred as an activity, the activity constitutes the combustion of the fuel. The notion of life is considered by Van Inwagen, not as a historical succession of social events in which an individual is involved, but rather in terms of whatever process (or processes) happens to be characteristic of all concrete biological organisms.[26]

[26]Though Van Inwagen leaves to biology the task of determining which specific processes are characteristic of concrete biological organisms, he points out that there are certain

Having specified what Van Inwagen means by 'activity constituting an event' and 'life', we are now in a good position to understand some of the consequences of Van Inwagen's proposed answer to the SCQ. One consequence of Van Inwagen's answer is the thesis that there are no inanimate material objects such as rocks, chairs or corpses. Although this is a rather counterintuitive consequence, Van Inwagen (1990: 105) argues that it does not contradict commonsense beliefs because assertions such as 'There are chairs' express, appearances notwithstanding, contentious philosophical theses which can be safely dispensed with insofar as there are no explanatory losses if we substitute them in the theories where they occur by assertions of the following type: 'There are certain material things that are arranged in chair-like fashion.'[27] In addition to the rejection of the existence of inanimate material objects, another consequence of Van Inwagen's proposed answer to the SCQ is the acceptance of cases of vague composition, i.e., cases in which it is indeterminate whether the $\mathcal{X}$s constitute a further entity y. These cases of vague composition arise because the notion of composition is defined in terms of the notion of part, as we earlier saw, and there are cases in which it is indeterminate whether a certain object c (e.g., an atom of carbon) is a part of some other object a (e.g., a woman named Alice), as the following story illustrates:

Alice drinks a cup of tea in which a lump of sugar has been dis-

general features that characterize a life. In particular, he (1990: 87-89) argues that a life is typically a well-individuated self-sustaining event and that a life is jealous in the sense that, if the activity of the $\mathcal{X}$s constitutes a life L at some time t, the activity of the $\mathcal{X}$s can only constitute L at t.

[27] I find this consequence to be unacceptable and I think that it gives us a very good reason to reject Van Inwagen's proposed answer, but I will not argue for this claim here. For some some discussion on this issue, see Hirsch (1993) and Rosenberg (1993).

> solved. A certain carbon atom that is part of that lump of sugar is carried along with the rest of the sugar by Alice's digestive system to the intestine. It passes through the intestinal wall and into the bloodstream, whence it is carried to the biceps muscle of Alice's left arm. There it is oxidized in several indirect stages (...) and is finally carried by Alice's circulatory system to her lungs and there breathed out as a part of a carbon dioxide molecule. (1990: 94)

As Van Inwagen points out, given that it is indeterminate whether the atom of carbon is a part of Alice at certain stages in the story, it is also indeterminate whether the activity of the $\mathcal{X}$s (which include the atom of carbon c and the material objects that constituted Alices's life immediately prior to her driking the tea) constitute a life at these same stages. Now, in addition to the fact that Van Inwagen's proposed answer sanctions the possibility of cases of vague parthood and vague composition, it also sanction the possibility of other instances of wordly indeterminacy. In particular — and this is the main reason that explains why Van Inwagen argues that Evans' argument is invalid, it sanctions the possibility of indeterminate identity.

In order to appreciate this, suppose that, instead of drinking tea and performing physical activity, Alice steps into the Cabinet where an experiment is performed on her such that it makes indeterminate whether the activity of the $\mathcal{X}$s that constitutes Alice's life immediately prior to her stepping into the Cabinet persists afterwards. After the experiment, if a person, which we dub 'Celia', steps out of the Cabinet, it appears that the identity statement 'Alice=Celia' is indeterminate in truth-value in virtue of some form of wordly

indeterminacy that is generated by the Cabinet. Consequently, since the intended conclusion of Evans' *reductio* contradicts one of the straightforward consequences of Van Inwagen's proposed answer to the SCQ, it is perfectly understandable why Van Inwagen argues that the *reductio* is invalid: blocking Evans' argument allows him to vindicate the plausibility of Life as an answer to the SCQ. Having clarified the background discussion that motivates Van Inwagen's resistance to the *reductio*, let us consider in detail the objection.

4.5.2 The objection: Van Inwagen's semantics as a tool to block the *reductio*

Van Inwagen is well aware, as I previously said, that he must block Evans' argument to uphold Life as an answer to the SCQ. Accordingly, he adopts the following strategy: he develops a semantics (or, to be more precise, a fragment of a semantics) that involves the basic formal machinery used in Evans' *reductio* and, having done that, he shows that the *reductio* turns out to be invalid in this semantics. The semantics he proposes is based on two key ideas:

> First, since certain sentences of the language we shall be treating are to be thought as neither true nor false, we shall need more than two truth-values. Secondly, if identity is indeed vague, then the semantic relation between name and thing named must also be vague. If, for example, 'Alpha' definitely names x and it is neither definitely true nor definitely false that $x{=}y$, then it

> seems inevitable to suppose that it is neither definitely true nor definitely false that 'Alpha' names y. (1990: 246)

Thus, the semantics that Van Inwagen proposes is a non-bivalent semantics that allows the possibility of cases of referential indeterminacy that stem from worldly indeterminacy.[28] Keeping this in mind, let me now introduce the basic vocabulary of the formal language (call it $\mathcal{L}$) in which Van Inwagen evaluates Evans' argument, the recursive rules for the construction of complex expressions in $\mathcal{L}$ and the semantic machinery that Van Inwagen uses to interpret the sentences of $\mathcal{L}$.

The basic vocabulary of $\mathcal{L}$ includes only the elements used in Evans' *reductio*: individual constants (a, b...), variables (x, y...), the operators 'λ', '$\neg$' and '$\bigtriangledown$' and the two-place identity predicate '='. Using this basic vocabulary, Van Inawagen presents the following recursive rules, which are used to construct more complex expressions in $\mathcal{L}$ such as predicates, lambda abstracts and sentences:

> Rules for predicate formation
>
> (P_1) Given any constant 'a' and any variable 'x', 'a=x' is a predicate.
>
> (P_2) Given any predicate 'Φ', '$\neg\Phi$' and '$\bigtriangledown\Phi$' are predicates too.
>
> (P_3) Nothing else is a predicate.
>
> Rules for lambda abstract formation

[28] I have argued previously in §4.4.3.2 that using a non-bivalent model to interpret Evans' argument is problematic since such models are unfaithful to the intentions of Evans, but I will not press Van Inwagen on this issue.

(A_1) Given any predicate 'Φ' and any variable 'x', '$\lambda\Phi$' is a lambda abstract if 'x' occurs in 'Φ'.

(A_2) Nothing else is a lambda abstract.

Rules for sentence formation

(S_1) Given any constants 'a' and 'b', 'a=b' is a sentence.

(S_2) Given any constant 'a' and any abstract '$\lambda\Phi$', '$a[\lambda\Phi]$' is a sentence.

(S_3) Given any sentence 'S', '$\neg$S' and '$\bigtriangledown$S' are sentences too.

(S_4) Nothing else is a sentence.

Once the basic vocabulary of $\mathcal{L}$ and the recursive rules for forming complex expressions are in $\mathcal{L}$ place, Van Inwagen introduces the semantic machinery he uses to interpret the sentences of $\mathcal{L}$, which is based on the notion of *model*.[29] Van Inwagen characterizes a model for $\mathcal{L}$ as a triple <U, P, i> where U is a non-empty set of objects, P is a set of pairs of objects belonging to U such that it is indeterminate whether the two members of each pair are identical with each other,[30] and i is an interpretation of the language $\mathcal{L}$ that assigns to each constant of $\mathcal{L}$ an object belonging to U, to each abstract and each predicate of $\mathcal{L}$ both an *extension* and a *frontier* and to each sentence of $\mathcal{L}$ a truth-value.[31] The notions of extension and frontier are characterized by Van Inwagen in the following fashion:

[29]Van Inwagen's notion of model differs slightly from the traditional one: whereas a model for a set of formulas Γ is typically conceived as an interpretation that makes all formulas of Γ true, Van Inwagen construes models as mere interpretations.

[30]When two objects are members of a pair that is a subset of P, they are said to be *paired* in P.

[31]The following notation will be used in the remainder of this section: given an interpretation i of $\mathcal{L}$, the referent of a constant 'α' will be denoted as '$i_r[\alpha]$', the extensions of

> We are to think intuitively of each abstract [and each predicate] as denoting (relatively to a model) a property; the extension of an abstract [or a predicate] is, intuitively, the class of objects that definitely have that property, and the frontier of the abstract [or a predicate] is the class of objects that neither definitely have nor definitely lack that property. (1990: 249)

Let me illustrate these two notions with an example. Suppose that U has as members only a and b (which are respectively denoted by 'α' and 'β' in a certain interpretation i of $\mathcal{L}$) and that they are such it is indeterminate whether 'α=β' is true. Since the basic vocabulary and the formation rules for complex expressions in $\mathcal{L}$ allow us to construct abstracts and predicates such as 'α=x', these predicates and abstracts turn out to have certain extensions and frontiers in U: in particular, $a \in i_e[\alpha{=}x]$ and $b \in i_f[\alpha{=}x]$.

After presenting the notion of model, Van Inwagen then proceeds to give a series of semantic rules that, given a certain model $M{=}<\text{U}, P, i>$ for the language $\mathcal{L}$, govern the attribution of extensions and frontiers to both abstracts and predicates:

> (SR_1) Given any identity predicate of the form 'α=x' in $\mathcal{L}$, we have $i_e[\alpha{=}x]{=}\{x{:}\ x{=}i_r[\alpha]\}$ and $i_f[\alpha{=}x]{=}\{x{:}\ <x, i_r[\alpha]>;\ x{\in}\text{U};\ i_r[\alpha]{\in}\text{U};\ \bigtriangledown(x{=}i_r[\alpha])\}$.
>
> (SR_2) For any predicate Φ, if $i_e[\Phi]{=}\text{E}$ and $i_f[\Phi]{=}\text{F}$, then $i_e[\neg\Phi]{=}\text{U-}(\text{E}\cup\text{F})$ and $i_f[\neg\Phi]{=}\text{F}$.

a predicate 'Φ' and of a lambda abstract '$\lambda\Phi$' will be denoted respectively as '$i_e[\Phi]$' and '$i_e[\lambda\Phi]$', the frontiers of a predicate 'Φ' and a lambda abstract '$\lambda\Phi$' will be denoted as '$i_f[\Phi]$' and '$i_f[\lambda\Phi]$ and the truth-value of a sentence S will be denoted as 'i_v[S]'.

(SR$_3$) For any predicate Φ, if $i_f[\Phi]$=F then $i_e[\bigtriangledown\Phi]$=F and $i_f[\bigtriangledown\Phi]=\emptyset$.

(SR$_4$) For any predicate Φ, if $i_e[\Phi]$=E and $i_f[\Phi]$=F then $i_e[\lambda\Phi]$=E and $i_f[\lambda\Phi]$=F

Having laid down these semantic rules for the attribution of extensions and frontiers to abstracts and predicates, Van Inwagen introduces a series of semantic rules to provide truth-values for sentences in a model M of $\mathcal{L}$. In the model that Van Inwagen proposes for $\mathcal{L}$, the sentences may have three truth-values: truth (which is denoted '1'), falsity (which is denoted '0') and neither truth nor falsity (which is denoted '$\frac{1}{2}$'). The semantic rules are the following:

(TV$_{1.1}$) $i_v[\alpha=\beta]$=1 iff something is the referent of both α and β.

(TV$_{1.2}$) $i_v[\alpha=\beta]=\frac{1}{2}$ iff nothing is the referent of both α and β and the referents of α and β are paired.

(TV$_{1.3}$) $i_v[\alpha=\beta]$=0 otherwise

(TV$_{2.1}$) $i_v[\lambda x[\Phi x]\alpha]$=1 iff the referent of α belongs to the extension of $\lambda x[\Phi x]$

(TV$_{2.2}$) $i_v[\lambda x[\Phi x]\alpha]=\frac{1}{2}$ iff the referent of α does not belong to the extension of $\lambda x[\Phi x]$ and either (i) the referent of α belongs to the frontier of $\lambda x[\Phi x]$ or (ii) the referent of α is paired in P with an object b such that b belongs to the extension or to the frontier of $\lambda x[\Phi x]$.

(TV$_{2.3}$) $i_v[[\lambda x[\Phi x]]\alpha]$=0 otherwise

($\text{TV}_{3.1}$) $i_v[\neg \text{S}]=1$ iff $i_v[\text{S}]=0$

($\text{TV}_{3.2}$) $i_v[\neg \text{S}]=0$ iff $i_v[\text{S}]=1$

($\text{TV}_{3.3}$) $i_v[\neg \text{S}]=\frac{1}{2}$ iff $i_v[\text{S}]=\frac{1}{2}$

($\text{TV}_{4.1}$) $i_v[\bigtriangledown \text{S}]=1$ iff $i_v[\text{S}]=\frac{1}{2}$

($\text{TV}_{4.2}$) $i_v[\bigtriangledown \text{S}]=0$ otherwise

Using these semantic rules for assigning truth-values for sentences in a model of $\mathcal{L}$ in conjunction with the previous rules for abstracts and predicates, Van Inwagen shows that that there is a model in which Evans' *reductio* turns out to be invalid (i.e., a model in which it fails to be '1-preserving'). To appreciate this, consider a very simple model M of $\mathcal{L}$ in which U contains just two elements —namely, a and b— which are respectively designated by 'a' and 'b' and which are also such that 'a=b' is neither true nor false. Thus, according to M, U=$\{a, b\}$, P=$<a, b>$ and i is such that $i_r[a]=a$, $i_r[b]=b$, $i_e[x=a]=\{a\}$, $i_f[x=a]=\{b\}$, $i_e[x=b]=\{b\}$ and $i_f[x=b]=\{a\}$.

If we evaluate Evans' argument using this model, we obtain the following results. (1) is true because $\{a, b\}$ is a subset of P. Thus, we can obtain $i_v[a=b]=\frac{1}{2}$ on the basis of $\text{TV}_{1.2}$ and then $i_v[\bigtriangledown(a=b)]=1$ using $\text{TV}_{4.1}$. (2) also comes out true in the model. We know that $i_e[\bigtriangledown(a=x)]=\{b\}$. Thus, in virtue of SR_4, we have $i_e[\lambda x[\bigtriangledown(a=x)]]=\{b\}$ and this yields, if we use $\text{TV}_{2.1}$, $i_v[[\lambda x[\bigtriangledown(a=x)]]b]=1$. (3) is true as well. It is clear that $i_v[a=a]=1$. This entails $i_v[\bigtriangledown(a=a)]=0$ in virtue of $\text{TV}_{4.2}$, which in turn allows us to obtain $i_v[\neg\bigtriangledown(a=a)]=1$ using $\text{TV}_{3.1}$. But (4) is false in the model, as the following reasoning shows. It is given that $i_f[a=x]=\{b\}$. Thus, we can obtain $i_e[\bigtriangledown(a=x)]=\{b\}$ using SR_3, which in turn allows us

to derive $i_e[[\lambda x[\bigtriangledown(x{=}a)]{=}\{b\}$ on the basis of SR_4. This enables us to get $i_v[\lambda x[\bigtriangledown(x{=}a)]a]{=}\frac{1}{2}$ using $TV_{2.2}$, which is equivalent to $i_v[\neg\lambda x[\bigtriangledown(x{=}a)]a]{=}\frac{1}{2}$ in virtue of $TV_{3.3}$. Thus, considering that the inference from (3) to (4) fails to be truth-preserving in this model, Van Inwagen concludes that Evans' *reductio* may be blocked.

4.5.3 Some difficulties with Van Inwagen's proposal

In this subsection, I want to cast light on certain difficulties that undermine Van Inwagen's response to the *reductio*. The first difficulty I want to focus on consists in the existence of a tension between one of the key ideas on which Van Inwagen's semantics is based and the way in which the semantics is implemented using the notion of model. Recall that one of the core ideas on which Van Inwagen's semantics is grounded is the existence of certain instances of referential indeterminacy that are due to worldly indeterminacy of identity. Let me briefly mention one more time the example that Van Inwagen uses to illustrate this idea:

> (E) If 'Alpha' definitely names x and it is neither definitely true nor definitely false that $x{=}y$, then it seems inevitable to suppose that it is neither definitely true nor definitely false that 'Alpha' names y.

Thus, according to Van Inwagen, instances of indeterminate identity that are due to worldly indeterminacy generate indeterminacy in the reference relation. Although this idea is meant to provide a way to block the *reductio*, it is important to notice that it is not properly implemented by the semantic

machinery that Van Inwagen introduces. Indeed, when Van Inwagen presents the notion of a model for $\mathcal{L}$, the notion of interpretation is taken to assign *determinately* to each constant of $\mathcal{L}$ an object in U. Thus, in order to secure the possibility of referential indeterminacy in the model that he presents, Van Inwagen does not appeal to the notion of interpretation (in which each constant is determinately attributed a referent) but rather to the notion of *fringe referent* (which is defined in terms of the set of pairs P) as the following passage shows:

> It will occasionally be useful to call the objects with which the referent of a constant is paired the *fringe referents* of that constant. The objects with which an object is paired are to be thought of as the objects such that it is indefinite whether that object is identical with them, and the *fringe referents* of a constant are to be thought as the objects such that it is indefinite whether that constant denotes them. (1990: 249)

As we can appreciate, the notion of fringe referent of a constant depends on the notion of (determinate) referent of a constant. Moreover, according to Van Inwagen, the referential indeterminacy displayed by a constant 'k' vis-à-vis the objects that are its fringe referents is due to the fact that it is indeterminate whether the referent of k is identical with any these objects. But, if it is indeed indeterminate whether the referent of 'k' is identical with any of the objects that belong to the set of fringe referents of 'k' (which amounts to say that the referent of 'k' and any of the members of the set of fringe referents of 'k' are *paired*), there is a clear conflict between this chacterization of the referential indeterminacy of 'k' stemming from the worldly

indeterminacy of identity between the referent of 'k' and the fringe referents of 'k' and the following characterization of paired objects that Van Inwagen provides when he defines P:

> A pairing P in a universe is a (possibly empty) set of two-membered (pairs) sets of members of that uinverse. These are to be 'genuinely' two-membered: $\{x, x\}[=\{x\}]$ cannot be a member of the pairing. If x and y, $x \neq y$, are members of a pair (belonging to a certain pairing), they are said to be paired (in that paring). (*ibid.*)

The tension now becomes patent. Van Inwagen aims to construct a semantics that captures the possibility of referential indeterminacy stemming from worldly indeterminacy of identity. To do this, he introduces the notion of model, which accounts for referential indeterminacy in terms of the notion of fringe reference. However, the notion of fringe reference depends on the notion of a pairing of objects, which in turn presupposes that these objects are, not indeterminately identical, but rather *distinct*. Consequently, considering that the semantics that Van Inwagen develops is intended to capture referential indeterminacy due to worldly indeterminacy of identity, it seems to be ill-suited for that purpose in light of the fact that the basic notions upon which the semantics is built (e.g., the notion of pairing) presuppose in their very definition a view according to which objects are either identical or distinct with each other, but not indeterminately identical.[32]

[32]Mena (MS) presents an argument slightly different from the one I give above (his argument involves considering the properties that Van Inwagen attributes to objects in the foundations of his semantics whereas mine focuses on the definitions of the basic

The second difficulty that I want to highlight consists in the fact that the semantics that Van Inwagen proposes generates a rather awkward result. This result is an interpretation of Evans' *reductio* in which both (2) and (4) are attributed value $\frac{1}{2}$ and (5) is attributed value 0. This interpretation is problematic because, according to Van Inwagen himself, it puts in question the plausibility of the semantics.[33] Thus, to defend his semantics, Van Inwagen presents an argument that aims to vindicate such an interpretation. Unfortunately, the argument founders, as I show below, in virtue of the fact that Van Inwagen does not take into account the existence of some *semantic relationships* that hold between the basic elements of (2), (4) and (5). Now, before explaining the nature of these semantic relationships, let me rehearse briefly how Van Inwagen shows that the semantics he proposes yields the abovementioned awkward result and how he answers the charge that this result makes the semantics implausible.

Suppose that we adopt the semantic apparatus that Van Inwagen proposes (i.e., the basic vocabulary of $\mathcal{L}$, the recursive rules for constructing complex expressions in $\mathcal{L}$, the notion of model and the semantic rules that provide an interpretation to all the elements of $\mathcal{L}$ in a model). Suppose further that, instead of considering the model M, we consider the model M* =<U*, P*, i*> in which U*=$\{a, c\}$, P*=<a, c> and i* is such that $i^*_r[a]=a$

notions underlying Van Inwagen's semantic machinery), but our diagnosis is the same: 'It seems like Van Inwagen's semantics is designed to deal with determinate objects, not with indeterminate ones.'

[33]The existence of a model that, given the semantics that Van Inwagen proposes, makes (2) and (4) indeterminate and (5) false casts doubt on the semantics because the inference from (2) and (4) to (5) is an instance of the argument schema that expresses, as Williamson (2002: 279) remarks, 'the canonical way in which we establish that objects are distinct', and cannot presumably be questioned.

and $i^*_r[b]=a$. Given these conditions, the application of the semantic apparatus to M* yields the following truth-values for (2), (4) and (5) respectively:

$i^*_v[\lambda x[\bigtriangledown(x=a)]b]=\frac{1}{2}$

$i^*_v[\neg\lambda x[\bigtriangledown(x=a)]a]=\frac{1}{2}$

$i^*_v[\neg(a=b)]=0$

Van Inwagen (1990: 252) admits that, even though this interpretation cannot be used to block Evans' *reductio* because the defenders of the *reductio* typically assume that the truth-value '$\bigtriangledown(a=b)$' is 1 (and, consequently, only models such as M in which one of the inferences is non-1-preserving allow us to invalidate Evans' argument), it does raise the following question: 'Does this fact tend to render our semantical fragment implausible?' To answer this question, Van Inwagen presents the following argument, which is aimed at defending the plausibility of his semantics:

> Consider a formal instance of the Nonidentity of Discernibles that contains only one constant, 'a'; let the abstract it contains be [the one used in (2) and (4)]:
>
> $\lambda x[\bigtriangledown(x=a)]a,\ \neg\lambda x[\bigtriangledown(x=a)]a \vdash \neg(a=a)$
>
> Any model that pairs the referent of 'a' with something will asign $\frac{1}{2}$ to each of the premises of this inference form, and any model at all will assign 0 to the conclusion. Perhaps it is implausible to assign $\frac{1}{2}$ to either of the premises. We have already considered that question. Given that we do, however, we should certainly want the conclusion to have the value 0. And should matters

> really be different if we assigned a second constant 'b' the same referent as 'a' and replaced the last occurrence of 'a' in the first premise and one of the occurrences of 'a' in the conclusion with 'b?' (1990: 264)

Van Inwagen presumably intends the question that he raises at the end of this passage to be rethorical. Unfortunately, he does not realize that, even if 'a' and 'b' are assumed to have the same referent, the two inferences

(I_1) $\lambda x[\bigtriangledown(x{=}a)]a,\ \neg\lambda x[\bigtriangledown(x{=}a)]a \vdash \neg(a{=}a)$

(I_2) $\lambda x[\bigtriangledown(x{=}a)]b,\ \neg\lambda x[\bigtriangledown(x{=}a)]a \vdash \neg(a{=}b)$

cannot be put on the same footing because the constants 'a' and 'b' respectively display in I_1 and I_2 different semantic relationships.

In order to show this, let me first illustrate the notion of semantic relationship by rehearsing a puzzle that Fine (2003: 606-608) calls the *antinomy of the variable* and showing how he solves it. Suppose that two variables x and y range over the same domain (say, the natural numbers). In light of this supposition, x and y seem to have the same semantic role in the expressions '$x{>}0$' and '$y{>}0$' since both expressions seem to be mere notational variants of the same mathematical statement. However, if we consider the expression '$x{>}y$', the semantic roles of x and y seem to be different since one could not substitute x by y (or *vice versa*) without expressing a different mathematical statement.[34] Thus, the puzzle arises because we feel compelled to agree

[34]Fine characterizes the notion of semantic role of a meaningful expression in terms of the purely linguistic or nonconventional aspect it has. This characterization is convenient insofar as it does not force us to choose any particular candidate.

with two seemingly incompatible claims: the variables x and y have the same semantic role but the pairs of variables x,x and x,y have different semantic roles.[35]

After considering several different of ways to handle the puzzle and showing that they all are unsatisfactory, Fine presents his own suggestion. In a nutshell, he argues that the puzzle may resolved if we reject the assumption that any difference in semantic role between two expressions has to be grounded in their intrinsic semantic features —an assumption he calls *semantic intrinsicalism*. The rejection of semantic intrinsicalism as a way to solve the puzzle is vividly expressed in the following passage:

> Once we have specified the range and the independence in value, we will then have a complete description of the semantic behavior of the variables; there is nothing more (at least at the extensional level) to be said about their role. But if this is so, then it is clear that the intrinsicalist doctrine, *no difference in semantic relationship without a difference in semantic feature*, must fail. For the intrinsic semantic features of any two variables will be the same —it will in effect be given by the specification of their range, whereas the intrinsic semantic features of the pairs x_1, x_2, say, and x_1, x_1, will be different, since the former will assume any pair of values from the given range while the latter will only

[35]Fine (2003: 607fn2) notices that this puzzle is, in a sense, parallel to Frege's puzzle: given that the names 'Cicero' and 'Tully' refer to the same object, 'Cicero' and 'Tully' seem to have the same semantic role in the statements 'Cicero is Roman' and 'Tully is Roman', but they also seem to have different semantic roles in the identity statements 'Cicero is Cicero' and 'Cicero is Tully'. This affinity between the puzzles is crucial for my argument, which extrapolates Fine's observations on semantic connections between variables to the case of constants.

> assume *identical* pairs of values. If we are merely informed of the intrinsic features of two variables, then we cannot tell whether they assume their values independently of one another (should they be distinct) or whether they always assume the same value (should they be the same). (2003: 622-623)

Let me flesh this out in more detail. Fine argues that, if we reject semantic intrinsicalism, we may consistently assert that the semantic role of x and y is the same and that the semantic roles of the pairs of variables x,x and x,y differ. This is due to the fact that, in the case of x and y, their semantic role exclusively depends on the range they have (which is the same) since each variable holds no semantic connections and that, in the case of x,x and x,y, the semantic roles of the pairs differ because there is a semantic relationship in the case of the first pair that forces the values adopted by the variables to always be the same. This semantic relationship is absent in the case of the second pair, as it is evidenced by the fact that the values that can be adopted by the variables are independent of each other and may consequently be distinct.

Now, what are the implications of Fine's remarks vis-à-vis Van Inwagen's defense of his semantics? Van Inwagen suggests that there is no crucial difference between I_1 and I_2 if we suppose that 'a' and 'b' have the same referent[36] and that, if we are willing to attribute certain truth-values to the premises and the conclusion of I_1 given the model M* using his semantics, we should be willing to attribute the same truth-values to the premises and

[36]There seems to be a serious problem here: how can Van Inwagen consistently assume that 'a' and 'b' have the same referent if he has previously granted that 'a=b' is indeterminate? However, I will not press Van Inwagen on this issue.

the conclusion of I_1. However, even if we suppose that 'a' and 'b' have the same referent, there is a crucial difference between I_1 and I_2: the pair of constants a, a that occurs in the two premises and the conclusion of I_1 exhibits a certain semantic relationship that guarantees that the value that the constants adopt is the same. This result is guaranteed in the case of I_2 by the supposition that 'a' and 'b' have the same referent, but *it does not flow from the semantic relationship on the pair a, a*. Thus, Van Inwagen cannot assume that, just because he supposes that the constants 'a' and 'b' have the same referent, the inferences I_1 and I_2 can be put automatically on the same footing. An advocate of Evans' *reductio* has room to argue that, even if it is unproblematic that Van Inwagen's semantics attribute value $\frac{1}{2}$ to the premises of I_1 and value 0 to its conclusion,[37] this unproblematic character does not surface in I_2 because it is due to the semantic relationship on the pair a, a that is absent in the case of I_2.

Let me recap. I have argued in this section that the objection that Van Inwagen addresses to Evans' *reductio* founders because of two reasons: (1) though Van Inwagen' semantics intends to provide a model for indeterminate identity, its basic notion (i.e., that of model) presupposes that objects are determinate and (2) Van Inwagen's response to the claim that his semantics has implausible consequences (i.e., casting doubt on the Non-Identity of Discernibles) fails because he does not realize that the pairs of constants a, a and a, b have different semantic relationships even if 'a' and 'b' are taken to have the same referent. In light of this, I conclude that Van Inwagen's

[37] I think it is deeply problematic that Van Inwagen's semantics attribute value $\frac{1}{2}$ to both premises of I_1, since this amounts to make the contradiction $\lambda x[\bigtriangledown(x{=}a)]a \wedge \neg\lambda x[\bigtriangledown(x{=}a)]a$ not entirely false. But I will not press Van Inwagen on this issue.

objection fails to undermine Evans' *reductio*.

4.6 Williams' response to the *reductio*

As I mentioned in §4.1, Williams (2008) has recently argued that there is an effective way to block Evans' *reductio* which does not appeal to the traditional strategies (e.g., arguing that the lambda abstract does not express a property or appealing to a non-bivalent logic). The key idea underlying Williams' proposal consists in exploiting a 'loophole' in Evans' *reductio*: since Evans does not rule out cases in which referential indeterminacy arises from worldly indeterminacy and referential indeterminacy provides us with an effective way to block the *reductio* (as we saw previously in §4.3), showing that there are cases in which worldly indeterminacy does give rise to referential indeterminacy allows us in principle to resist the *reductio*. To assess Williams's response, I present in §4.6.1 the loophole that he attempts to exploit in a more detailed way, highlighting the key ideas that he relies on to illustrate its nature. Having done this, I rehearse in §4.6.2 the framework that Williams introduces to exploit the loophole and establish the existence of indeterminate identity statements that stem from a form of worldly indeterminacy. Finally, in §4.6.3, I argue that Williams' proposal to resist the *reductio* founders because the kind of worldly indeterminacy that he considers is unable to generate the kind of semantic indeterminacy required to block Evans' *reductio*.

4.6.1 The loophole in Evans' *reductio*

After laying down in the first two sections of his paper the conditions that an effective response to Evans' *reductio* ought to fulfill according to him and highlighting the import of the indeterminate identity thesis for both ethical and metaphysical debates, Williams argues that there is a loophole in the *reductio* that can be exploited to block it. In order to show the loophole, Williams (2008: 145) first presents a couple of theses widely endorsed. The first one is that, though we can attempt to block Evans' argument (in particular, the property abstraction and the contrapositive of property concretion steps) by arguing that the lambda abstracts do not refer to genuine properties, this objection can be resisted successfully by adopting a deflationary attitude vis-à-vis properties. The second thesis is that to put up any resistance against the property abstraction and the contrapositive of property concretion steps in Evans' *reductio* after adopting a deflationary view of properties amounts to miss Evans' point.

Though Williams agrees wholeheartedly with the first thesis, he only agrees partially with the second. To be more specific, Williams argues that the second thesis involves one correct thought and one wrong thought. The correct thought, which he endorses, is captured by the following sub-thesis:

> (C) Unless 'a' and 'b' are referentially indetermiante, then both the following are true:
>
> • if it is indeterminate whether a is identical with b, then a has the property of being indeterminately identical to b.
>
> • if a has the property of being indeterminately identical to b,

then it is indeterminate whether a is identical with b.

The wrong thought, which he rejects, is captured by the following subthesis:

> (W) If 'a' is referentially indeterminate, then it is so in virtue of *semantic indecision*.

What are the consequences of Williams' endorsement of C and his rejection of W? Williams' acceptance of C has a rather straightforward implication: there is no way to block the property abstraction or contrapositive of property concretion steps in Evans' *reductio* other than diagnosing referential indeterminacy in 'a' or 'b'. Thus, given that the restrictions on a successful response to Evans' *reductio* that Williams lists at the onset of his paper bar us from putting into question the initial assumption (1), the premise (3) or the conclusion (5), the only way for Williams to resist Evans' *reductio* is *via* referential indeterminacy.

But isn't this simply conceding Evans' point? After all, granting that Evans aims to show that no identity statements are indeterminate in truth-value in virtue of worldly indeterminacy, it seems that admitting that the only way to block the *reductio* is through a diagnosis of referential indeterminacy amounts to agree with Evans that identity statements cannot be indeterminate in truth-value in virtue of worldly indeterminacy. This is precisely where the loophole in Evans' *reductio* lies: although referential indeterminacy arises in many circumstances, not all these circumstances involve *semantic indecision*.

Cases of referential indeterminacy due to semantic indecision are well-known. For instance, the expression 'gavagai' uttered by a hypothetical native in Quine's discussion about radical translation exhibits referential indeterminacy in virtue of *semantic indecision* because rabbits, undetached rabbit-parts and temporal-slices of rabbits are all determinate entities for Quine. What generates referential indeterminacy in this case is that, though all the candidate referents for the expression 'gavagai' are determinate objects, the expression does not attach to one in particular but rather *hovers* over all them. Now, in addition to 'gavagai', several other terms are taken to display referential indeterminacy in virtue of semantic indecision in the literature.[38] But semantic indecision is not the only source of referential indeterminacy for Williams. In some cases, the source seems to lie, not in some limitation of our language, but rather in how the world is. To show this, Williams first introduces the following considerations concerning the reference relation and the potential sources of reference failure:

> The reference relation is the joint upshot of what we do to fix the meaning of our words and how the world is. A theory of reference must deal with both, as is shown by reference failure. For the sake of argument, suppose that one of the things we have to do in order to refer to a thing is to acquire the capacity to recognize that thing again perceptually under a range of circumstances. Some attempts to introduce a name referring to an object might then fail, not because the world is not co-operating, but because

[38] In particular, Field (1973) has persuasively argued that 'mass' and many other scientifc terms are referentially indeterminate due to semantic indecision.

> we have not done *our* part. (...) On the other hand, there are cases where we seem to have done our part correctly, but where we fail to refer to anything because the world is not co-operating, as with Macbeth's demonstrative *this dagger*, for example. (2008: 146-147)

Having introduced these considerations, Williams then argues that, just as reference failure may have two distinct sources —one grounded on the language and the other grounded on the world— in virtue of the structure of the reference relation, referential indeterminacy may also have two different sources: one grounded on the language, which corresponds to semantic indecision, and one grounded on the world, which corresponds to worldly indeterminacy.

Consequently, the loophole in Evans' *reductio* consists in the fact that, in principle, there can be instances of referential indeterminacy that arise in virtue of worldly indeterminacy. Now, in order to successfully exploit the loophole to block Evans' *reductio*, one must show that there are *actual* cases of referential indeterminacy that stem from worldly indeterminacy and that these cases of referential indeterminacy give rise to identity statements that are indeterminate in truth-value. Williams attempts to show precisely this using a modal framework that I describe in the following section.

4.6.2 Williams' modal framework: multiple actualities

The framework that Williams presents to exploit the aforementioned loophole and block Evans' *reductio* is based on a certain view (or, more accurately, a family of views) about the nature of possible worlds. According

to all these views, the actual world is fundamentally distinct in kind from possible worlds. Whereas the actual world is typically regarded as some kind of concrete physical entity that encompasses us and all our surroundings according to these views, possible worlds are traditionally considered as mere unrealized ways things could have been.

In addition to the thesis that the actual world is different in kind from possible worlds, another central tenet of the family of views that Williams relies on to develop his framework is that, among the myriad of possible worlds that represent the unactualized possibilities, there is a privileged member that represents the only possibility that is realized —i.e., the actual world. This privileged member, which is often referred as the *actualized world* to distinguish it both from the actual world and from the rest of possible worlds,[39] admits different ways of being construed, which in turn depend on the different conceptions of possible worlds as representing unactualized possibilities that are endorsed by philosophers. For instance, those who construe possible worlds as maximally consistent sets of propositions (e.g., Adams 1974) consider the actualized world as the maximally consistent set of propositions that provides an accurate and complete description of the actual world whereas those who construe possible worlds as maximal properties that the actual world could have instantiated (e.g., Forrest 1986) consider the actualized world as the maximal property that happens to be instantiated by the actual world.

After stressing (i) the distinction in kind between the actual world and the merely possible worlds and (ii) the existence of an actualized world that

[39]For further discussion of the distinction between the actual world and the actualized world in the context of the ersatzist program, see Lewis (1986: 136-142).

corresponds to or represents the actual world as underlying his framework, Williams remarks that, since the nature of the correspondence relation between the actual world (which he calls *reality*) and the actualized world (which he calls *actuality*) is not completely determined, this leaves open a certain gap between the thesis that we inhabit a single reality and the thesis that there is only one actuality that corresponds to this reality that can be used to provide a model that explains how worldly indeterminacy surfaces. Here is more precisely Williams illustrates the occurrence of the gap by considering distinct ways to construe the notion of an actualized world that corresponds to the actual world:

> *Prima facie*, many distinct world-properties could be instantiated by reality; many of the images may depict reality equally well. On a *sui generis* approach to correspondence, why not think that *sui generis* relation is many-one. The idea that there is a single [actualized] world corresponding to reality now needs argument; I contend that no general argument for this conclusion is available. (2008: 149)

Thus, since the correspondence relation between reality and actuality is not determinate insofar as multiple candidates may fulfill equally well the role of actuality, we can use this to provide a model of how worldly indeterminacy surfaces by maintaining that, if reality is somehow indeterminate, this worldly indeterminacy manifests itself in the fact that it is indeterminate which actuality corresponds to reality. This framework of *multiple actualities* is quite versatile, as Williams shows, because it enables us to account for distinct types of worldly indeterminacy, such as indeterminacy in existence

or indeterminacy in location. For instance, granting that it is indeterminate whether some object o is present at some location l, the framework allow us to account for this indeterminacy in terms of the existence of an actuality at which o is located at l and an actuality where o is located elsewhere.

Moreover, the adoption of the framework of multiple actualities is defended by Williams, not only because it enables us to account for many different types of worldly indeterminacy, but also on the grounds that it provides an explanation of the conditions under which a statement is true if the actual world displays indeterminacy. In a nutshell, the idea is that a statement is true at the actual world iff it is true relative to all the relevant actualities. To illustrate this, let us suppose that that there is a rock dubbed 'Rock' at the base of Mount Vesuvius such that it is worldly indeterminate whether it is part of the volcano or not. Notwithstanding this worldly indeterminacy, we can affirm that the statement 'Rock lies somewhere in Italy' is true if, as it is presumably the case, Mount Vesuvius and its immediate surroundings are located in Italy in all the relevant actualities.

With the framework in place, Williams then shows how it can be used to generate a case of indeterminate identity to block Evans' *reductio*. The strategy that he embraces is the following: rather than claiming that identity statements of indeterminate truth-value arise in cases where two objects are distinct at one actuality and distinct at another (a move that Williams rejects since it sanctions failures of the transitivity of identity), he presents a fission scenario in which the framework of multiple actualities is applied to a case of worldly indeterminacy stemming from a survival puzzle.

Here is the fission scenario. Williams considers a certain amoeba called

'Sue' that splits into two distinct amoebas respectively called 'Sally', which wanders off to the west after the fission, and 'Sandy', which wanders off to the east. He also supposes that the fission event is such that (i) Sue survives past the split, (ii) it is indeterminate whether Sue survives as Sally or Sandy and (iii) this indeterminacy is not semantic but rather worldly. Since the survival of Sue as either Sally or Sandy is, *ex hypothesi*, a case of worldly indeterminacy, we can apply to it the framework of multiple actualities. If we do this, the worldly indeterminacy of the fission scenario can be accounted in terms of the existence of two distinct actualities that represent equally well the actual world:

> Actuality$_1$: The amoeba named 'Sue' survives as the amoeba named 'Sally' and the amoeba named 'Sandy' is the newly minted amoeba.
>
> Actuality$_2$: The amoeba named 'Sue' survives as the amoeba named 'Sandy' and the amoeba named 'Sally' is the newly minted amoeba.

The application of the framework of multiple actualities to the fission case shows, according to Williams, that the two names 'Sally' and 'Sandy' are referentially indeterminate since 'Sally' names the surviving amoeba in Actuality$_1$ and the newly minted one in Actuality$_2$ whereas 'Sandy' names the newly minted amoeba in Actuality$_1$ and the surviving amoeba in Actuality$_2$. But where does the referential indeterminacy that permeates both names arise from? Williams answers as follows:

> Not from any failure in semantic conventions. Metaphysically, I am supposing that it is indeterminate where the surviving amoeba Sue is after the fission. Since I have introduced the names 'Sally' and 'Sandy' (in part) by pointing to an amoeba at a certain location, this ontic indeterminacy *induces* referential indeterminacy. The (ontically based) referential indeterminacy produces a vague identity statement. 'Sue=Sandy' is true at one actual[ized] world, but false at the other. So, overall, it is indeterminate in status. (2008: 151)

Thus, since the worldly indeterminacy concerning the location of the surviving amoeba Sue after the fission induces a form of referential indeterminacy for both 'Sally' and 'Sandy' and this referential indeterminacy gives rise in turn to the identity statement 'Sue=Sally' that is indeterminate in truth-value, Williams seems to have the means to block effectively Evans' argument. Indeed, as we saw previously in §4.3, diagnosing referential indeterminacy in either the constant 'a' or the constant 'b' allows us to block the inference from (3) to (4) in the *reductio*. Accordingly, if one substitutes in Evans' argument 'a' by 'Sandy' and 'b' by 'Sue', one can block that inference and, given that the referential indeterminacy of 'Sandy' stems from a form of worldly indeterminacy (as the framework of multiple actualities shows), it seems that the way in which the inference is blocked undermines the specific interpretation of the *reductio* that Evans intended.

4.6.3 Some difficulties with Williams' response

I turn now to a critical assessment of Williams' response. My goal in this subsection consists in showing that it does not undermine Evans' *reductio*. To do this, I present in the first place a couple of minor problems that respectively arise vis-à-vis the way in which Williams characterizes the loophole in Evans' *reductio* and vis-à-vis the way in which Williams defends the framework of multiple actualities. Having done that, I introduce later what I take to be the main difficulty with Williams's proposal, which is consists in the fact that the kind of wordly indeterminacy that Williams considers is unable to generate the kind of referential indeterminacy needed to exploit effectively the loophole.

Let me present the first minor problem with Williams' response. In order to illustrate the existence of the loophole, Williams makes, as we previously saw in §4.6.1, a series of remarks about the reference relation and cases of referential failure that he later extrapolates to cases of referential indeterminacy. According to him, the reference relation is the joint upshot of two distinct factors: our linguistic practices and how the world is. In light of this two-factor dependence, Williams maintains that, just as cases of referential failure may arise in virtue of a linguistic fact (e.g., a shortcoming of our linguistic practices) or a worldly fact (e.g., the non-existence of an object o), cases of referential indeterminacy may also be the result of either a linguistic fact or a worldly fact.

Even though nobody disputes the claim that reference is the joint upshot of the two aforementioned factors, someone might have qualms with Williams' assumption that the contribution of each factor to the reference

relation can be distinguished clearly from the other —an assumption which he trades on to argue that cases of referential indeterminacy may arise either from semantic indecision or from worldly indeterminacy. Indeed, this assumption can be criticized from a Quinean perspective as the product of the two dogmas of empiricism. To be more specific, someone who agrees with the case that Quine mounts against the two dogmas might raise the following considerations about reference, which echo Quine's (1951: 39) remarks on truth and science:

> It is nonsense, and the root of much nonsense, to speak of the linguistic component and the factual component of a referring expression. Taken collectively, reference has its double dependence upon linguistic practice and world, but this duality is not significantly traceable into the words of a language taken one by one.

Thus, in order to establish conclusively the existence of the loophole (which depends on the existence of a clear distinction between linguistic factors and worldly factors in the reference relation), Williams must find a way to quell these Quinean-inspired worries. Now, considering that I am not pessimistic vis-à-vis the prospects of distinguishing clearly the linguistic and worldly components that make up the reference relation,[40] I am content with

[40]Kripke's causal picture of reference (1980: 91-92) provides us with a basic account of how we may distinguish the worldly component of reference from the linguistic component (at least in the case of names): the linguistic component is given by the convention adopted during a baptism to use a certain name to refer to a certain individual and by the practice adopted by subsequent users of the name to refer to that individual when using the name whereas the wordly component is given by the individual. Of course, this basic account has to be modified to take into account certain problems such as the difficulty to explain

just mentioning this possible objection against Williams without pushing him on this issue.

Here is the second minor problem. After Williams shows the existence of the loophole in Evans' argument, he develops the framework of multiple actualities to exploit it. As we previously saw in §4.6.2, this framework is very attractive in virtue of many features. In particular, it allows us to present a unified account of how different types of worldly indeterminacy surface and it enables us to provide an account of the general conditions under which a statement is true at the actual world if the actual world is itself indeterminate.

But the framework is also problematic in one respect: Williams defends its plausibility by claiming that there is no general argument for the conclusion that there is a single actualized world that corresponds to the actual world. However, this claim trades on an assumption that he makes —namely, that the actual world is somehow indeterminate and that possible worlds are determinate entities. Now, if this assumption is questioned, an argument for the conclusion that there is only a single actualized world that corresponds to the actual world becomes available.

This argument is grounded on a view presented by Hartshorne (1965: 190), the core of which is that 'definiteness is the soul of actuality'. On this view, our world is actual because it is a determinate entity and merely possible worlds, which are unrealized ways our world could have been, are *determinable* entities in the sense that they admit a certain degree of vari-

cases, such the one discussed by Evans (1973: 196), in which the subsequent linguistic practice of speakers deviates from the convention initially adopted at the baptism. In future work, I intend to develop a more refined version of Kripke's proposal that will both deal with Evans' cases and be applicable to referring expressions other than names.

ation with respect to certain distinct features.[41] To illustrate this, consider a possible world that corresponds to the possibility that I had a brother instead of a sister. On the view that Harsthorne presents, this possible world is a determinable entity because it admits a degree of variation with respect to certain features (e.g., the place of birth of my brother, his date and time of birth, the method by which he was delivered, etc.) that provide different specifications for the possibility whereas the actual world admits no such variation.

Granting that all the merely possible worlds are determinables and that only the actual world is determinate, it seems that no possible world can correspond or represent completely and in full detail the actual world. Indeed, in order to represent completely and in full detail the actual world (i.e., in order to be an actualized world), any merely possible world would have to be as determinate as the actual world iteself. But only one world is as determinate as the actual world on the view that Harsthorne presents —namely, the actual world itself. In light of this, the view that Harsthorne presents allows us to provide an argument for the thesis there is a single actualized world that corresponds to the actual world by relying on the fact that there is only one world determinate enough to provide an accurate and complete representation of the actual world.[42]

[41]The notion of a determinable entity can be illustrated by considering the property *being human*. Indeed, this property is a determinable entity because its many instances vary with respect to a number of features (e.g., age, sex, marital status, etc.) that provide specifications of the property in terms of more discriminating properties (e.g., *being a 33-year-old married woman*). For a recent discussion of the notion of a determinable entity, see Funkhouser (2006).

[42]Lewis (1986: 198) points out that the simplest way that a possible world could represent the actual world is by identity. Now, it is important to notice that, even though the thesis that representation works by identity faces serious difficulties when we consider

Even though there is, *pace* Williams, an argument available for the conclusion that there is a single actualized world that corresponds to the actual world, I will not push him on this issue because the framework of multiple actualities that he develops is grounded on an ersatzist view of possible worlds and versions of ersatzism typically presuppose a distinction between the actual world and the actualized world. Thus, I am also content with mentioning this potential objection to Williams's response without endorsing it.

Let me turn now to what I consider to be the main difficulty with Williams' response. Williams maintains that, by exploiting the loophole through the use of the framework of multiple actualities, one may effectively block Evans' *reductio*. As we previously saw, the strategy that he relies on consists in considering a fission scenario that is supposed to illustrate a case of referential indeterminacy in an identity statement (thus making it indeterminate in truth-value) stemming from worldly indeterminacy. What I aim to show is that, even if we grant that Williams' fission scenario exemplifies both worldly indeterminacy and referential indeterminacy, the kind of referential indeterminacy that is required in order to block Evans' *reductio* is not generated by the kind of worldly indeterminacy that he introduces.

To appreciate this clearly, it will be rather useful to rehearse in detail a series of distinctions between distinct types of referential indeterminacy and worldly indeterminacy that Eklund has established. After pointing out that he takes the expression 'our thoughts and practices' to encompass whtaever it is that fixes the reference of our terms (which including our interactions with

the issue of how a possible world can represent a non-actualized possibility of the actual world (specifically, it faces the problem of accidental intrinsics), the thesis seems to face no major difficulties when we consider the issue of how the actualized world represents the actual world.

the environment in order to accomodate the truth of externalism), Eklund draws the following distinctions:

> Let an expression be *semantically indeterminate with respect to meaning* (*SM-indeterminate*) iff, for all that is determined by our thoughts and practices, there could be several candidate-referents of the expression. An expression is *semantically indeterminate with respect to reference* (*SR-indeterminate*) iff there are actually several candidates for the expression. Let an expression be *ontologically indeterminate with respect to meaning* (*OM-indeterminate*) iff, provided that there are indeterminate objects, the expression refers to an indeterminate object (...) An expression is *ontologically indeterminate with respect to reference* (*OR-indeterminate*) iff it actually refers to an indeteminate object. (2008: 128)

With these distinctions laid out, let us now consider how they can be used to critically discuss Williams's response. We saw previously in §4.6.1 that, according to Williams, the only effective way to block the *reductio* is by diagnosing referential indeterminacy in either 'a' or 'b'. Let us grant then that 'a' is referentially indeterminate, i.e., that there is no fact of the matter regarding what the reference of 'a' is. If there is no fact of the matter regarding what the reference of 'a' is, this is tantamount to affirm that there *are* several candidate-referents for 'a', in which case 'a' is SR-indeterminate rather than SM-indeterminate. But the SR-indeterminacy of 'a' is not enough by itself to effectively block Evans' *reductio*: it must be such that it stems from some form of worldly indeterminacy. In light of this,

it seems that the only way to guarantee that the SR-indeterminacy of 'a' blocks effectively Evans' *reductio* consists, given Eklund's distinctions, in 'a' being OR-indeterminate also rather than OM-indeterminate.

Does Williams' fission scenario correspond to this model? We know that the names 'Sally' and 'Sandy' are SR-indeterminate since both admit two candidate-referents —namely, the surviving amoeba and the newly minted amoeba. Now, in order to show that Evans' *reductio* does not go through in the fission scenario where 'Sue' stands for 'b' and 'Sally' stands for 'a', it has to be the case that 'Sally' is SR-indeterminate in virtue of being OR-indeterminate. *But this not what occurs.* Indeed, as Williams himself acknowledges in a passage previously quoted, the worldly indeterminacy in the fission scenario is due to the fact that, metaphysically, he assumes that it is indeterminate where the surviving amoeba Sue is after the fission. Thus, 'Sue' is OR-indeterminate because it is indeterminate where the surviving amoeba is located, but this worldly indeterminacy is clearly unsuited to make the SR-indeterminacy of 'Sally' block Evans' *reductio* —only the OR-indeterminacy of 'Sally' could do the job in this case. But 'Sally' exhibits no OR-indeterminacy since the amoeba that it refers to is located both at Actuality_1 and Actuality_2 in the west after the fission —i.e., its location is determinate. Considering all this, I conclude that Williams' response to Evans *reductio* ultimately fails because, even if we overlook other contentious issues, the kind of worldly indeterminacy that he introduces in his fission scenario cannot be used to obtain the kind of referential of referential indeterminacy that he relies on to block Evans' argument.

4.7 Barnes' objection

In this section, I present and provide a critical assesment of another recent response to Evans' *reductio*, which is developed by Barnes (2009). The core of Barnes' suggestion to block Evans' *reductio* consists in arguing that there are some rather strong motivations that support a *counterpart-theoretic* interpretation of Evans' argument and that, given this interpretation, one may diagnose an equivocation in the argument that invalidates one of the inferences. To provide a thorough analysis of Barnes' response to the *reductio*, I go over briefly in §4.7.1 the framework of modal vagueness that she assumes as background for her discussion of Evans' argument —a framework which is motivated by certain similarities between certain modal operators and the indeterminacy operator. Having done this, I introduce in §4.7.2 the motivations that Barnes gives to offer a counterpart-theoretic interpretation of Evans' argument within the framework of modal vagueness she adopts and I show how, given this interpretation, Evans' argument does fail in virtue of an equivocation. Finally, in §4.7.3, after examining closely the motivations that Barnes presents for the counterpart-theoretic interpretation of the argument, I reject them and I offer two additional arguments that cast doubt on Barnes' proposal.

4.7.1 Barnes' framework of modal vagueness

After rehearsing Evans' *reductio* at the onset of her paper, Barnes presents the framework of modal vagueness in which much of her subsequent discussion is embedded. This framework, which is similar in several respects to the

view propounded by Akiba (2000), is grounded on the existence of certain parallels between a semantics that is traditionally adopted for modal statements (call it PW-semantics) and a semantics that is adopted for determinate and indeterminate statements (call it P-semantics).

According to PW-semantics, the modal operators involved in modal statements have to be interpreted as quantifiers ranging over possible worlds, and this enables us to provide truth-conditions for modal statements. For instance, it is traditionally claimed that statements of the form '$\Box$S' are true iff S is true at every possible world and that statements of the form '$\Diamond$S' are true iff S is true at some possible world. Now, according to P-semantics,[43] the determinacy and indeterminacy operators (which are respectively symbolized by '$\triangle$' and '$\triangledown$') have to be interpreted as quantifiers ranging over not possible worlds, but rather *precisifications*, and this enables us to provide truth-conditions for sentences involving the determinacy and indeterminacy operators. Specifically, statements of the form '$\triangle$S' are taken to be true iff S is true at all the admissible precisifications and statements of the form '$\triangledown$S' are taken to be true iff S is true at some admissible precisification and false at another.[44]

The abovementioned structural similarities between the PW-semantics for modal operators and the P-semantics for the determinacy/indeterminacy operators sanctions a *pseudo-modal* treatment of indeterminacy (*pseudo-modal*

[43]P-semantics is typically endorsed by authors who embrace supervaluationist accounts of indeterminacy. In addition to Akiba (2000), other authors who adopt a form of P-semantics include Fine (1975) and McGee and Laughlin (1994).

[44]In light of this, the modal operator that bears the closest similarity to the indeterminacy operator '$\triangledown$' is not the possibility operator '$\Diamond$' but rather the contingency operator '$\mathcal{C}$' because a statement is contingent iff it is true at some possible world and false at another.

because the precisifications are *like* possible worlds). But Barnes maintains that one should go further: instead of considering the precisifications over which the determinacy and indeterminacy operators range as merely similar to possible worlds, it is convenient to treat these precisifications as a subset of the set of possible worlds.

This move, which entails no longer a pseudo-modal but rather a *modal* treatment of indeterminacy, is motivated by several considerations, as Barnes remarks. The first is that it allows us to provide an account of the phenomenon of indeterminacy by providing an account of modality because the notion of determinacy turns out to be, as Barnes (2009: 84) writes, 'a form of restricted modality'. A second consideration put forward to defend the move is that it is seemingly well adapted to provide an account of cases of worldly indeterminacy when used in conjunction with an ersatzist view of possible worlds. Indeed, given that most ersatzist views of possible worlds endorse a distinction between the actual world and the actualized world, the framework that Barnes assumes allows us to account for the existence of indeterminacy in the actual world in terms of the existence of several distinct candidates to play the role of the actualized world. Any one of these candidates, which are all precisifications of the actual world, respresents itself correctly as being the actualized world but, when they are viewed in their capacity of surrogates of the actual world, they all turn out to be precise misrepresentations of an indeterminate object. Consequently, a key advantage of the framework is that it allows us to cash out instances of worldly indeterminacy in terms of the existence of multitude of inaccurate representations of the actual word.

Further considerations that Barnes mentions in favor of the framework of

modal vagueness she assumes include the fact that (i) it respects both the principle of bivalence and that of excluded middle (i.e., within the framework, every statement is either true or false at any precisification and, for any statement S and any precisification P, either S or its negation is true at P), that (ii) it it is appealing in terms of ontological economy and (iii) that it enables us to capture the intuitive thesis that, if the world is indeterminate, there is no precise representation of it that is accurate.

After highlighting the abovementioned considerations in favor of the framework of modal vagueness, Barnes (2009: 85) remarks that 'this type of modal representation of indeterminacy is in no way essential to [my] response to Evans' argument' because, even though the framework provides some support to the counterpart-theoretic interpretation, the latter can be endorsed independently of the former. Though I do not agree with this claim, I will withhold my objections until §3.7.3 and turn to a brief exposition in the following section of the counterpart-theoretic interpretation of Evans' *reductio*.

4.7.2 The counterpart-theoretic interpretation of Evans' *reductio*: motivations and implications

Before presenting the counterpart-theoretic interpretation of Evans' *reductio* that defended by Barnes, let me rehearse briefly an observation that she makes regarding the existence of different ways of being a counterpart-theorist. One way of being a countepart-theorist consists is adopting the basic primitives, the postulates and the translation schema that Lewis (1968) proposed for sentences of quantified modal logic. But one can be also a counterpart-theorist in another way —namely, one can admit that the truth-

conditions of modal predications such as 'a could have been F' are given not in terms of strict transworld identity (i.e., a is F in some possible world W) but rather in terms of some similarity relation (i.e., some entity in W that bears a counterpart relation to a is F) and that the similarity relation in question is not constant but varies depending on several contextual factors. Barnes holds that she is a counterpart-theorist only in the second way.[45]

Having made clear the way in which she embraces counterpart-theory, Barnes then presents the main motivations to defend a counterpart-theoretic intepretation of Evans' *reductio*. The first motivation consists in the fact that the way in which modal predications are interpreted in counterpart theory appears to be particularly well suited to provide an account of the semantic framework of modal vagueness. Indeed, since the framework of modal vagueness provides truth-conditions for statements governed by the operators '$\triangle$' and '$\triangledown$' in terms of precisifications as we saw in §3.7.1 and the precisifications in question are not the original statements themselves (but rather interpretations of the statements in which indeterminacy is eliminated), a counterpart-theoretic interpretation of '$\triangle$' and '$\triangledown$' seems to capture adequately the idea that the truth-conditions of, say, '$\triangledown$S', are given in terms of something that is very much like S in many respects (e.g., a statement S* that provides a precise interpretation for all the vague terms that appear in S), but that is not really identical to it.

Similarly, if we consider how the framework of modal vagueness deals with

[45]In light of this, it appears that Barnes is committed to the thesis that counterpart theory offers *analyses* of both determinate and indeterminate statements. But this makes her view vulnerable to some objections that Fara and Williamson (2005) raise against the thesis that counterpart theory provides correct analyses of modal statements. For further discussion, see §3.7.3 below.

cases of worldly indeterminacy, it seems that counterpart-theory allows us to capture neatly the key idea underlying the treatment of worldly indeterminacy in the framework. Recall that, as we saw previously in §4.7.1, cases of worldly indeterminacy are handled within the framework by appealing to the distinction between the actual world and the actualized world in the following way: if it is worldly indeterminate whether some object o has F, there are multiple precisifications (at least one that represents o having F and one that represents o not having F) that are candidates to represent the actual world, none of which is the actual world itself. Counterpart theory enables us to capture this thesis in the following terms: if it is worldly indeterminate whether o has F, then there is a precisification of the actual world at which a counterpart of o has F and a precisification of the actual world at which a counterpart of o lacks F. Barnes sums up nicely the first motivation in favor of the counterpart-theoretic intepretation in the following passage:

> Precisifications give us a helpful model for distinguishing determinacy and indeterminacy, but they are not the way things are —in reality things (words, objects, properties, etc.—depends on your theory of vagueness) are vague. With this in mind, then, it's best not to speak of identity across precisifications. Rather, we have actual (vague) things and then we have their precisified counterparts. We can thus latch on to true claims about actual things based on the way things are in precisifications, while at the same time maintaining that nothing here in the vague world is identical with anything in the precisifications. (2009: 87-88)

In addition to the fact that counterpart theory enables us to provide a

neat model of the relation between a vague entity —whether it is is an object, a word or something else— and its precisifications (in the sense that, just a precisification of a vague entity is similar but not identical with the entity it stands for, a counterpart is similar but not identical with the entity it bears the counterpart relation to), the second main motivation that Barnes mentions for adopting a counterpart-theoretic interpretation of Evans' *reductio* is that, just as counterpart-theory allow us to explain the context-sensitivity of modal statements, it also allows us to account for the context-sensitivity of determinate and indeterminate statements alike.

Let me illustrate this with some examples. It is well established that modal statements are context-sensitive to the extent that their truth-value varies when they are uttered in different contexts. For instance, 'Some birds could have been mammals' is true when it is uttered at a context set given by the laws of logic, but it is false when uttered at a context given by the laws of biology. Now, counterpart theory allows us to explains rather nicely this context-senstivity in the following way: whereas the context given by the logical laws determines a certain counterpart relation C such that there are mammals at some other worlds that bear C to actual birds, the context given by the biological laws determines another counterpart relation C* such that no mammal at another world bears C* to any actual bird.

Barnes maintains that both determinate and indeterminate statements also exhibit context-sensitivity. For instance, if we consider the statement 'The sun is determinately hot', the statement turns out to be true if it is uttered at a context that only involves the objects of our solar system but it is false it it is uttered at a context that involves all the stars of our galaxy.

This context-sensitivity of both determinate and indeterminate statements is a trait that counterpart theory allows us to account for. Indeed, just as we may explain using counterpart theory why a shift in the context in which a modal statement occurs causes a shift in the set of relevant worlds used to provide truth-conditions for the statement,[46] we may explain using also counterpart theory why determinate and indeterminate statements are subject to the same type of shifts, as Barnes points in the following passage:

> 'x is determinately F' is true in a context just in case the counterpart relation invoked in that context only picks as counterparts of x things that are F. And the same utterance that is true in one context can be false in another context, simply because a different counterpart relation is invoked, one which picks out some x-counterparts that are not F. The counterpart-theoretic notion of the salience of a particular counterpart relation tells us *why* there is a shift in what precisifications are admissible. (2009: 88)

Summing up, there are two main motivations according to Barnes to adopt a counterpart-theoretic interpretation of Evans' argument: (i) this interpretation appears to be particularly well suited to account for the framework of modal vagueness, which many authors use to provide semantics for indeterminate statements such as the ones used in Evans' *reductio* and (ii) this interpretation enables us to explain why both determinate and indeterminate statements exhibit context-sensitivity.

[46]Lewis (1971) offers a classical illustration of this type of shifts using the following example: even if we grant that I am identical to my body, the sentences 'I could have been distinct from my body' and 'My body could have been distinct from my body' are not equivalent since 'I' selects the *personal* counterpart relation and 'my body' selects the *bodily* counterpart relation.

Having laid down these two motivations to adopt the counterpart-theoretic interpretation, Barnes proceeds to show the main implication of the interpretation —i.e., how it allows us to block Evans' *reductio*. As I previously mentioned, the general strategy that Barnes follows consists in arguing that, once we have accepted a counterpart-theoretic interpretation of the *reductio*, we may diagnose an equivocation that blocks one of the inferences. Here is a bit more precisely how Barnes argues for the existence of the equivocation.

Barnes initially reminds us that, since the way in which things are referred to often determines for counterpart theorists which standard of similarity is salient (and, thus, which counterpart relation has to be selected in a given occasion), they typically appeal to this fact to argue that, even if the object denoted by 'a' and the object denoted by 'b' are identical at the actual world, the identity statement 'a=b' is not necessarily true. Indeed, given that counterpart theorists maintain that the truth-conditions of modal predications must be given, not in terms of the objects mentioned in the predications but rather in terms of their counterparts, the identity statement 'a=b' is necessarily true iff every counterpart of a is also a counterpart of b. But, since different contexts select different counterpart relations (in particular, the use of 'a' selects the a-counterpart relation and the use of 'b' selects the b-counterpart relation), a counterpart theorist may argue that 'a=b' is not necessarily true even if 'a' and 'b' denote the same object in the actual world because there are some a-counterparts of this object in other possible worlds that are not b-counterparts (and *viceversa*).

Barnes mentions that this feature of counterpart theory may be used to block the inference from (i) $\Diamond\neg(a{=}b)$ and (ii) $\neg\Diamond\neg(a{=}a)$ to the conclusion

that a and b have distinct properties by diagnosing an equivocation in the argument. Even if it is the case that a could not have been distinct from a and that a could have been distinct from b, these two claims are true, for a counterpart theorist, relative to different contexts that select distinct counterpart relations, but there is no single context in which both claims are true at once. A counterpart-theoretic translation of both (i) and (ii) allows us to appreciate this:

(i_t) Some a-counterparts of a are not b-counterparts of it.

(ii_t) All a-counterparts of a are a-counterparts of it.

As the translations clearly illustrate, (i) is true because, relative to a context that selects the b-counterpart relation, the salient counterparts of a are b-counterparts and some of these are in principle not a-counterparts of a and (ii) is true because, relative to a context that selects the a-counterpart relation, the salient counterparts of a are a-counterparts and all of these are trivially a-counterparts of a. Thus, considering that (i) and (ii) are true but only relative to different contexts (which are respectively given by the senses of 'b' and 'a'), it is then patent, that on a counterpart-theoretic interpretation of modal operators, one cannot infer from (i) and (ii) the conclusion that a and b are distinct without switiching from one context to another —i.e., without switching counterpart relations and, consequently, equivocating.

Now, just as the abovementioned feature of counterpart theory enables us to block the inference from (i) and (ii) to the conclusion that a and b are distinct by diagnosing an equivocation, Barnes argues that we may also appeal to this feature to block the inference from the steps (2) $\lambda x[\bigtriangledown(x{=}a)]b$

and (4) $\neg\lambda x[\bigtriangledown(x{=}a)]a$ to the conclusion (5) $\neg(a{=}b)$ in Evans' *reductio*. The strategy is the following: if we agree with Barnes that the indeterminacy operator '$\bigtriangledown$' that occurs in Evans' *reductio* is a kind of modal operator and that a counterpart-theoretic interpretation of Evans' *reductio* is justified in virtue of the two motivations that Barnes puts forward, one may argue that both (2) and (4) are true but that, since each step is true relative to a different context, the conclusion (5) does not follow without switching contexts at some point —which is tantamount to equivocating. Having presented the way in which Barnes' blocks the *reductio*, let me provide now a critical assessment of it.

4.7.3 A critical assessment of Barnes' response

I turn in the present subsection to a critical review of Barnes' response to the *reductio*. I proceed as follows: I consider first the motivations that Barnes gives in favor of the counterpart-theoretic interpretation and, after providing solid reasons to reject both of them, I present two arguments against the counterpart-theoretic interpretation.

Let us consider once more Barnes' first motivation for adopting a counterpart-theoretic interpretation of Evans' *reductio*, which I recount using her own words (2009: 86): 'counterpart theory lends it self particularly well to the semantic framework in question because of its unique treatment of transworld identity'. Barnes also maintains, as we previously saw, that the framework is not essential to the response she gives to the *reductio* because one can accept the counterpart-theoretic construal of the indeterminacy operator and nevertheless maintain a non-modal account of prescisifications. However, this

remark, in conjunction with other observations that she makes, seriously undercuts the first motivation.

Let me explain this in detail. The feature of counterpart theory that Barnes relies on to motivate the counterpart-theoretic interpretation of Evans' *reductio* is 'its unique treatment of transword identity'. I take this to mean that the main reason to invoke counterpart theory is that, since transworld identity is such a strict and unyielding relation, counterpart theory enables us to make sense of the intuition that things might have been distinct even though they are in fact identical.

In addition to the fact that counterpart theory allows us to vindicate many intuitions about the occurrence of contingent identity in a series of puzzles, Stalnaker (1986) mentions that there are two other motivations to be a counterpart-theorist: (i) Lewisian modal realism and (ii) anti-haecceitism. Now, since Barnes does not embrace either modal realism or anti-haecceitism, I take it that her main motivation to endorse counterpart theory is the fact that it enables us to vindicate intuitions about contingent identity. Considering this, Barnes (2009: 93) clearly undermines her main reason to be a counterpart theorist when she acknowledges that 'one might easily maintain that modality and indeterminacy are only limitedly analogous, and thus adopt counterpart theory for the determinacy operators while retaining a standard treatment of the modal operators.' Indeed, if one separates the counterpart-theoretic treatment of the indeterminacy operator in Evans' *reductio* from the counterpart-theoretic treatment of the modal operators and rejects the latter while endorsing the former, one eliminates a funadamental reason to be a counterpart theorist. Moreover, if one is neither a modal re-

alist nor an anti-haecceitist, then one has seems to have no motivations left to be a counterpart-theorist in the first place.

Let us consider the second motivation for the counterpart-theoretic interpretation, which consists in the fact the counterpart theory enables us to account for the context-sensitivity of both determinate and indeterminate statements. Now, I agree with Barnes that both determinate and indeterminate statements are context-sensitive and that counterpart theory enables us to account for this context-sensitivity. *But there are better ways to account for this feature*. Let me explain. The deployment of counterpart theory to account for 'hypothetical' statements (i.e., statements that 'point beyond themselves' such as modal or tensed statements) goes hand in hand with an attempt to analyze them in terms of other statements that are 'categorical' (i.e., statements that are 'self-contained' such as non-modal or tenseless statements).[47] Thus, anybody who maintains that counterpart theory allows us to provide a satisfactory account of certain 'hypothetical' statements seems to be committed to the claim that there is an effective *analysis* of these statements in terms of some other 'categorical' statements.

However, several authors have claimed that such an analysis is simply unavailable in many cases because there are truth-value discrepancies between the *analysantes* and the *analysanda*. For instance, Fara and Williamson (2005: 7-13) have argued that counterpart theory fails to provide adequate analyses of certain modal statements (in particular, modal statements in-

[47]This is particularly clear in the case of Lewis (1986), but authors who reject modal realism while embracing counterpart theory for modal (or tensed) statements also are in the business of providing analyses of these statements. For an illustration, see both Heller (1998) and Sider (2001).

volving the actuality operator ACT).[48] To show this, they first introduce the following modal statement (which is clearly unsatisfiable):

> (D) There could have been an object o such that it is actually F if and only if it is actually not-F.

After introducing D, they argue that it cannot be given a proper counterpart-theoretic analysis in any plausible construal of the actuality operator. Indeed, if we provide a counterpart-theoretic interpretation of the actuality operator as an existential quantifier ranging over actual counterparts, D comes out as:

> (D_1) There is some object o at some non-actual world such that it has some counterpart at the actual world that is F if and only if it has some counterpart at the actual world that is not-F.

And, if we construe the actuality operator as a universal quantifier ranging over actual counterparts, D comes out as:

> (D_2) There is some object o at some non-actual world such that all of its counterparts at the actual world are F if and only if all of its counterparts at the actual world are not-F.

Having presented these two options, Fara and Williamson show that neither D_1 nor D_2 provides an adequate analysis of D because, in models of counterpart theory in which o has no counterparts whatsoever at the actual

[48] Several authors have argued that a quantified modal logic expressive enough to capture the vast range range of modal judgments we endorse must include, in addition to the possibility operator '$\Diamond$' and the necessity operator '$\Box$', the actuality operator 'ACT', which is read as 'It is actual that...'. In particular, see Hazen (1976) and Forbes (1985) .

world, both D_1 and D_2 turn out to be satisfiable even if D is not. Considering this, since a successful deployment of counterpart theory on certain statements seems to commit us to the existence of a good counterpart-theoretic analysis of these statements and since there are statements (in particular, modal statements involving the actuality operator such as D) that cannot be given a good counterpart-theoretic analysis, it seems that that embracing counterpart theory entails some risks and potential costs.

Granting that the adoption of counterpart-theory may create some problems, we are then entitled to ask the following question: is there a way to account for the context-sensitivity of determinate and indeterminate statements without appealing to counterpart theory? The interest-relative theory of vagueness defended by Graff (2000) constitutes an attractive candidate to undertake this task. According to Graff's, the context-sensitivity of determinate and indetermiante statements can be accounted for in terms of the diversity of our interests in different occasions. To illustrate this, suppose that I happen to be standing next to a bucket filled with springwater. For Graff, the statement 'The bucket contains determinately a lot of water' may be true in some contexts and false in others depending on what I intend to do with the water. For instance, if I intend to quench my thirst, the statement is true in the context determined by this intention but, if I intend to irrigate thirty acres, the statement is false in the context determined by this intention.

What motivations are there to prefer Graff's theory of vagueness over counterpart theory as a way to account for the context-sensitivity of determinate and indeterminate statements? There two main reasons. The first

one consists in the fact that, since Graff's theory was specifically developed to deal with indeterminacy and does not extend to modality, it is completely insulated from the kind of problem raised by Fara and Williamson for modal counterpart theory, a version of which could potentially arise for Barnes' proposal. The second motivation is that Graff's theory enables us to provide a much more refined account of the context-sensitivity of determinate and indeterminate statements than Barnes' proposal. Indeed, for Barnes, the context-sensitivity of a statement such as 'The sun is determinately hot' clearly depends on several factors, but her account does not provide a systematic way to identify them. Using Graff's theory, on the contrary, we may explain the context-sensitivity of the statement in terms of different factors that are well identified. Indeed, given that Graff's theory construes the predicate 'is hot' as *has significantly to A a higher temperature than is the norm N for C*, Graff's theory allows us to explain the context-sensitivity of the predicate in terms of (at least) three distinct factors:

> —the agent A for whom the assessment of temperature is significant
>
> —the norm N that is used to assess the termperature
>
> —the comaprison class C with respect to which the tempreature is evaluated

Thus, since Graff's interest-relative theory of vagueness provides us with a potentially less problematic and more refined account of the context sensitivity of determinate and indeterminate statements than counterpart theory, we have then solid grounds to reject Barnes' second motivation for her proposal.

Having critically examined the two main motivations that Barnes presents to defend a counterpart-theoretic interpretation of Evans' *reductio*, let me offer now two arguments against the interpretation. The first argument consists in showing that the counterpart-theoretic interpretation creates a serious distortion of Evans' original intentions. Let me explain this. Evans initially characterizes at the onset of his *reductio* 'a' and 'b' as singular terms —i.e., as terms that are such that each one purports to refer to a single object.[49] In light of this, Evans presumably intends that, when both 'a' and 'b' occur in the different steps of the *reductio*, each one is used to refer to a single object —presumably, a and b. Unfortunately, interpreting the *reductio* in counterpart-theoretic terms distorts this intention because, as Barnes (2009: 87) herself concedes, a person who interprets '$\bigtriangledown Fa$' in terms of the precisified counterparts of a does not really refer to a because 'strictly speaking, she's talking about a different word/object when dealing with a precisification.' In fact, she is talking about many different precisified objects that are located at different precisifications. Thus, since the counterpart-theoretic interpretation of the *reductio* forces us to treat the singular terms 'a' and 'b' that Evans introduces in a way that is contrary to his intentions, Barnes' proposal constitutes a non-charitable interpretation of the *reductio*.

The second argument consists in showing that the counterpart-theoretic interpretation of the *reductio*, which aims to provide a way to block Evans' *reductio* and thus vindicate the thesis that there are indeterminate identities, is problematic because it does not really achieve the goal that Barnes (2009:

[49]There are many different characterizations of singular terms. I follow here Strawson (1950: 320) who maintains that a singular term is an expression used to refer to 'some individual person or single object or particular event or place or process.'

94) sets for it —namely, to provide 'a way of understanding indeterminate identity that is both fully coherent and compatible with both identity and non-identity.' Indeed, in order to provide a way of understanding indeterminate identity, the counterpart-theoretic interpretation should in principle give us a clear explanation of indeterminate identity. Unfortunately, there is a persuasive consideration presented by Salmon in his discussion of Kripke's remarks about indeterminate identity that casts a serious doubt on the availability of such an explanation. To appreciate this, let us briefly remind that Kripke (1980: 51fn18) suggests in a footnote that counterpart theory might have some application to alleged cases of indeterminate identity. The primary motivation for this suggestion seems to be that, if identity is indeed indeterminate, the best way to conduct a semantic investigation of it is to appeal to counterpart theory to the extent that counterpart relations, which are surrogates of identity in the metalanguage, are vague and intransitive. But, even though this counterpart-theoretic interpretation may be semantically useful, it faces a serious shortcoming, which consists in the fact that it fails to provide a precise way of understanding what indeterminate identity is, as Salmon rightly complains in the following passage:

> [But] Kripke is proposing that an allegedly vague concept, identity, be investigated in terms of another vague concept, *counterparthood*. It is difficult to see how there is anything to be gained representing one vague concept by another. If the problem is that we lack a logic of vagueness, we can no more treat the former than the latter. If our purpose is to investigate the logic of identity (...), surely we are better off sticking to genuine identity, and do-

> ing the best we can, than than turning our attention elsewhere only to find the same obstacles arise there. (2005b: 305)

Putting Salmon's worry in slightly different terms, even though counterpart theory may be rather interesting taken as *semantic machinery* that aims to supply truth-conditions to indeterminate identity statements, it ultimately fails to deliver a genuine *way of understanding* indeterminate identity since such an a thing would require clarifying the allegedly vague notion (i.e., identity) in terms of certain precise notions, as is done in other cases.[50] Thus, since the notion of counterpart is an inherently vague notion, it is ill-suited to provide us with an effective way to understand the notion of identity on the assumption that this notion is also vague. This consideration casts a serious doubt on Barnes' project of using counterpart theory to block Evans' *reductio* and vindicate indeterminate identity.

4.8 Conclusion

Let me now recap. I have argued in this chapter that the thesis that identity may be indeterminate in certain cases in virtue of some form of worldly indeterminacy is incorrect by offering a systematic defense of Evans' argument. In light of this defense, I believe it is reasonable to conclude that, even if there are many things or concepts that exhibit indeterminacy, identity is not among them: if x and y are identical, x and y are also determinately identical. But this conclusion raises a question: what should we do with the puzzles and thought experiments mentioned at the onset of this chapter?

[50] I have in mind the vague notion of computability, which is typically clarified in terms of the precise notion of recursive function.

I want to suggest briefly here a way to handle them, which is consistent with the conclusion of Evans' argument: rather than locating the indeterminacy in the identity relation, a far more plausible way to handle these puzzles and thought experiments would be to locate the presence of indeterminacy in other concepts. Just to give a very succinct illustration of how the strategy would work, we could grant, if we consider the case of Van Inwagen's Cabinet, that there is indeed no fact of the matter regarding whether Alice is the same person as Celia but also argue that the indeterminacy present in this case is due to the fact that concept *person* exhibits some indeterminacy. A more refined and systematic development of this strategy is something I intend to focus on in future works.

Bibliography

[1] Akiba, K. (2000) 'Vagueness as a Modality' *Philosophical Quarterly* 50 (200): 359-370.

[2] Akiba, K. (2002) 'A Deflationist Approach to Indeterminacy and Vagueness' *Philosophical Studies* 107: 69-86.

[3] Almog, J. (1981) 'DThis and DThat: Indexicality Goes beyond That' *Philosophical Studies* 39 (4): 347-381

[4] Ayer, A. J. (1946) *Language, Truth and Logic*.

[5] Adams, R. M. (1974) 'Theories of Actuality' *Noûs* 8 (3): 211-231.

[6] Adams, R. M. (1979) 'Primitive Thisness and Primitive Identity' *Journal of Philosophy* 76 (1): 5-26.

[7] Ayer, A. J. (1946) *Language, Truth and Logic* 2^{nd} ed. London: Gollancz.

[8] Bach, K. (1994) *Thought and Reference*. New York: Oxford University Press.

[9] Baker, L. R. (1999) 'Unity without Identity: a new look at Material Constitution' *Midwest Studies in Philosophy* 23: 144-165.

[10] Baker, L. R. (2002) 'On making Things up: Constitution and its Critics' *Philosophical Topics: Identity and Individuation* 30: 31-51.

[11] Barnes, E. (2009) 'Indeterminacy, Identity and Counterparts: Evans reconsidered' *Synthese* 168: 81-96.

[12] Baxter, D. (1988) 'Identity in the Loose and Popular Sense' *Mind* 97: 575-582.

[13] Bealer, G. (2002) 'Modal Epistemology and the Rationalist Renaissance' in T. Szabó Gendler and J. Hawthorne (eds.) *Conceivability and Possibility* (New York: Oxford University Press), 71-125.

[14] Bealer, G. (2004) 'The Origins of Modal Error' *Dialectica* 58 (1): 11-42.

[15] Black, M. (1952) 'The Identity of Indiscernibles' *Mind* 61 (242): 153-164.

[16] Blanchette, P. (1999) 'Relative Identity and Cardinality' *Canadian Journal of Philosophy* 29 (2): 205-223.

[17] Boghossian, P. (2006) *Fear of Knowledge*. New York: Oxford University Press.

[18] Broome, J. (1984) 'Indefiniteness in Identity' *Analysis* 44 (1): 6-12.

[19] Brower, J. (2009) 'Medieval Theories of Relations' *Stanford Encyclopedia of Philosophy* at http://plato.stanford.edu/entries/relations-medieval

[20] Burrough, P. A. and Frank, A. U (1996) *Geographic Objects with Indeterminate Boundaries*. London: Taylor and Francis.

[21] Campbell, J. (2002) *Reference and Consciousness*. Oxford: Clarendon Press.

[22] Carnap, R. (1947) *Meaning and Necessity*. Chicago: University of Chicago Press.

[23] Carnap, R. (1950) 'Empiricism, Semantics and Ontology' *Revue Internationale de Philosophie* 4: 20-40.

[24] Cartwright, R. (1954) 'Ontology and the Theory of Meaning' *Philosophy of Science* 21 (4): 316-325.

[25] Cartwright, R. (1967) 'Classes and Attributes' *Noûs* 1 (3): 231-241.

[26] Cartwight, R. (1987) 'On the Logical Problem of the Trinity' in *Philosophical Essays* (pp. 187-200). Cambridge, MA: MIT Press.

[27] Church, A. (1943) Review of Quine's 'Notes on Existence and Necessity' *Journal of Symbolic Logic* 8: 45-47.

[28] Church, A. (1958) 'Ontological Commitment' *Journal of Philosophy* 55 (23): 1008-1014.

[29] Davidson, D. (1970) ''Events as particulars' *Noûs* 4 (1): 25-32.

[30] Della Rocca, M. (2002) 'Essentialism versus Essentialism' in T. Szabó Gendler and J. Hawthorne (eds.) *Conceivability and Possibility* (New York: Oxford University Press), 223-252.

[31] Devitt, M (2006) 'Intuitions in Linguistics' *British Journal for the Philosophy of Science* 57 (3): 481-513.

[32] Dummett, M. (1958) 'Truth' *Proceedings of the Aristotelian Society*: 141-162

[33] Dummett, M. (1981) *Frege. Philosophy of Language*. (2nd ed.) Cambridge: Harvard University Press

[34] Dummett, M. (1996) *The Seas of Language*. New York: Oxford University Press.

[35] Devitt, M. (2006) 'Intuitions in linguistics' *British Journal for the Philosophy of Science* 57: 481-513.

[36] Eklund, M (2008) 'Deconstructing Ontological Vagueness' *Canadian Journal of Philosophy* 38 (1): 117-140.

[37] Evans, G. (1978) 'Can there be vague objects?' *Analysis* 38 : 208.

[38] Evans, G. (1982) *The Varieties of Reference*. Oxford: Oxford University Press.

[39] Fara, M. and Williamson, T. (2005) 'Counterparts and Actuality' *Mind* 114 (453): 1-30.

[40] Feldman, F. (1969) 'Geach and Relative Identity' *Review of Metaphysics* 22 (3): 547555.

[41] Feldman, F. (1973) Feldman (1973) 'Sortal Predicates' *Noûs* 7 268-282.

[42] Feyerabend, P. (2010) *Against Method*. (4^{th} ed.) New York: Verso.

[43] Field, H. (1973) 'Theory Change and the Indeterminacy of Reference' *Journal of Philosophy* 70 (14): 464-481.

[44] Fine, K. (1975) 'Vagueness, Truth and Logic' *Synthese* 30 (3/4): 265-300.

[45] Fine, K. (2003) *Journal of Philosophy* 100 (12): 605-631.

[46] Fitch, G. W. (2004) *Saul Kripke*. Montréal: McGill's University Press

[47] Fodor, J. (1986) 'Why paramecia don't have mental representations' *Midwest Studies in Philosophy* 10: 3-23

[48] Forbes, G. (1985) *The Metaphysics of Modality*. Oxford: Oxford University Press.

[49] Forrest, P. (1984) 'Ways worlds could be' *Australasian Journal of Philosophy* 64: 15-24.

[50] Frege, F. (1884) *Die Grundlagen der Arithmetik*. Breslau: W. Koebner. Translated as *The Foundations of Arithmetic* by J.L. Austin (Oxford: Blackwell), second revised edition, 1974.

[51] Frege, G. (1892a) 'Über Begriff und Gegenstand' *Vierteljahrsschrift für wissenschaftliche Philosophie* 16: 192-205. Translated by P. Geach and

reprinted in *Collected papers on Mathematics, Logic and Philosophy* (Oxford: Basil Blackwell), 182-194.

[52] Frege, G. (1892b) 'Über Sinn und Betdeutung' *Zeitschrift für Philosophie und Philosophische Kritik*' 100: 25-50. Translated by M. Black and reprinted in *Collected papers on Mathematics, Logic and Philosophy* (Oxford: Basil Blackwell), 157-177.

[53] Frege, G. (1894) Review of E. G. Husserl's *Philosophie der Arithmetik Zeitschrift für Philosophie und Philosophische Kritik*' 100: 313-332. Translated by H. Kaal and reprinted in *Collected papers on Mathematics, Logic and Philosophy* (Oxford: Basil Blackwell), 195-209.

[54] Frege, G. (1979) *Posthumous Writings* (tr. Long and White). Chicago: The University of Chicago Press.

[55] Frege, G. (1980) *Philosophical and Mathematical Correspondence* (tr. H. Kaal). Chicago: The University of Chicago Press.

[56] French, S. and Krause, D. (2003) 'Quantum Vagueness' *Erkenntnis* 59: 97-124.

[57] Funkhouser, E. (2006) 'The Determinable-Determinate Distinction' *Noûs* 40 (3): 548-569.

[58] Garrett, B. (1991) 'Vague Identity and Vague Objects' *Noûs* 25 (3): 341-351.

[59] Geach, P. (1957) *Mental Acts*. London: Routledge and Kegan Paul.

[60] Geach, P. (1962) *Reference and Generality* 1^{st} ed. Ithaca: Cornell University Press. Third ed. (1980)

[61] Geach, P. (1967) 'Identity' *Review of Metaphysics* 21 (1): 3-12.

[62] Geach, P. (1969) 'A Reply' *Review of Metaphysics* 22 (3): 556559.

[63] Geach, P. (1973) 'Ontological Relativity and Relative Identity' in M. K. Munitz (ed.), *Logic and Ontology* (New York: New York University Press), 267-302.

[64] Gibbard, Alan (1975) 'Contingent Identity' *Journal of Philosophical Logic* 4 (2): 187-221.

[65] Goodman, N. (1972) 'Seven Strictures on Similarity' in N. Goodman (ed.) *Problems and Projects* (Indianapolis and New York: Bobbs-Merrill), 437-447.

[66] Goodman, N. (1977) *The Structure of Appearance* (3rd. ed.) Springer:

[67] Graff, D. (2000) 'Shifting Sands: an Interest-Relative Theory of Vagueness' *Philosophical Topics* 28 (1): 45-81

[68] Grandy, R. (2007) 'Sortals' in *Stanford Encyclopedia of Philosophy* at http://plato.stanford.edu/entries/sortals/

[69] Griffin, N. (1977) *Relative Identity*. Oxford: Clarendon Press.

[70] Häggqvist, Søren (2006) 'Essentialism and Rigidity' *Philosophical Quarterly* 56 (1): 275-283.

[71] Hanson, N. R. (1958) *Patterns of Discovery*. Cambridge: Cambridge University Press.

[72] Harman, G. (1977) *The Nature of Morality: An Introduction to Ethics*. New York: Oxford University Press.

[73] Hawley, K. (2002) 'Vagueness and Existence' *Proceedings of the Aristotelian Society* 102: 125-140

[74] Hawthorne, J. (2006) *Metaphysical Essays*. New York: Oxford University Press.

[75] Hazen, A. (1976) 'Expressive Completeness in Modal Languages' *Journal of Philosophical Logic* 5 (1): 25-46.

[76] Heck, R. G. (1998) 'That there might be Vague Objects (so fas as concerns Logic)' *The Monist* 81 (2): 274-296.

[77] Heck, R. G. (2000) 'Cardinality, Counting and Equinumerosity' *Notre Dame Journal of Formal Logic* 41 (3): 187-209.

[78] Heller, M. (1998) 'Counterpart Properties in Ersatz Worlds' *Journal of Philosophy* 95 (6): 293-316.

[79] Hirsch, E. (1993) 'Peter Van Inwagen's *Material Beings*' *Philsophy and Phenomenological Research* 53 (3): 687-691.

[80] Hirsch, E. (1997) 'Basic Objects: a reply to Xu' *Mind and Language* 12 (3-4): 406-412.

[81] Hyde, D. (1994) 'Why Higher-Order Vagueness is a Pseudo-Problem' *Mind* 103 (409): 35-41.

[82] Johnsen, B. (1989) 'Is vague identity incoherent?' *Analysis* 49: 103113.

[83] Johnston, M. (1987) 'Is there a Problem about Persistence?' *Proceedings of the Aristotelian Society, Supplementary Volumes* 61: 107-135.

[84] Johnston, M. (2005) 'Constitution' in F. Jackson and M. Smith (eds.) *The Oxford Handbook of Contemporary Philosophy* (New York: Oxford University Press), 636-677.

[85] Jubien, M. (1972) 'The Intensionality of Ontological Commitment' *Noûs* 6 (4): 378-387.

[86] Jubien, M. (1996) 'The Myth of Identity Conditions' *Philosophical Perspectives. Metaphysics* 10: 343-356.

[87] Kaplan, D. (1975) 'How to Russell a Frege-Church' *Journal of Philosophy* 72 (19): 716-729.

[88] Kaplan D. (1989a) 'Demonstratives' in J. Almog, J. Perry and H. Wettstein (eds.) *Themes from Kaplan* (New York: Oxford University Press), 481-563.

[89] Kaplan D. (1989b) 'Afterthoughts' in J. Almog, J. Perry and H. Wettstein (eds.) *Themes from Kaplan* (New York: Oxford University Press), 565-614.

[90] Keefe, Rosanna (1995) 'Contingent Identity and Vague Identity' *Analysis* 55 (3): 183-190.

[91] Kim, J. (1993) *Supervenience and Mind*. New York: Cambridge University Press.

[92] Kripke, S. (1963) 'Semantical Analysis of Modal Logic I' *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 9: 67-96

[93] Kripke, S. (1979) 'A Puzzle about Belief' in A. Margalit (ed.) Meaning and Use (Dordrecht: Dordrecht Reidel Publishing Co.), 239-283.

[94] Kripke, S. (1980) *Naming and Necessity*. Cambridge, MA: Harvard University Press.

[95] Kripke, S. (ms) *Time and Identity*. Unpublished lectures delivered at Cornell University in 1978.

[96] Lewis, D. K. (1968) 'Counterpart theory and quantified modal logic' *Journal of Philosophy* 65 (5): 113-126.

[97] Lewis, D. K. (1971) 'Counterparts of Persons and their Bodies' *Journal of Philosophy* 68 (7): 203-211

[98] Lewis, D. K. (1983a) 'New Work for a Theory of Universals' *Australasian Journal of Philosophy* 61 (4): 343-377.

[99] Lewis, D. K. (1983b) *Philosophical Papers. Vol. I* New York: Oxford University Press.

[100] Lewis, D. K. (1986) *On the Plurality of Worlds*. Oxford: Blackwell.

[101] Lewis, D. K. (1988) 'Vague Identity: Evans misunderstood' *Analysis* 48 (3): 128-130.

[102] Locke, J. (1995) *An Essay on Human Understanding*. Amherst, NY: Prometheus Books.

[103] Lowe, E. J. (1989a) *Kinds of Being*. Oxford: Blackwell.

[104] Lowe, E. J. (1989b) 'What is a Criterion of Identity?' textslPhilosophical Quarterly 39: 1-21.

[105] Lowe, E. J. (1993) 'Are the Natural Numbers Individuals or Sortals' *Analysis* 53 (3): 142-146.

[106] Lowe, E. J. (1994) 'Vague Identity and Quantum Indeterminacy' *Analysis* 54 (2): 110-114.

[107] Lowe, E. J. (2007) 'Sortals and the Individuation of Objects' *Mind and Language* 22 (5): 514-533.

[108] Machina, K. (1976) 'Truth, Belief and Vagueness' *Journal of Philosophical Logic* 5: 47-78.

[109] Maddy, P. (1990) *Realism in Mathematics*. New York: Oxford University Press

[110] Marcus, Ruth Barcan (1947) 'The Identity of Individuals in a Strict Functional Calculus of Second Order' *Journal of Symbolic Logic* 12 (1): 12-15.

[111] Markosian, N. (1998) 'Brutal Composition' *Philosophical Studies* 92: 211-249.

[112] Martí, G. (2003) 'The Question of Rigidity in New Theories of Reference' *Noûs* 37 (1): 161-179.

[113] Maudlin, T. (2003) 'Distilling Metaphysics from Quantuum Physics' in M. Loux and D. Zimmerman (eds.) *The Oxford Handbook of Metaphysics* (New York: Ocford University Press), 461-490.

[114] McGee, V. and McLaughlin, B. (1995) 'Distinctions Without a Difference' *Southern Journal of Philosophy* 33: 203-51.

[115] McGinn, C. (1981) 'Modal Reality' in R. Healey (ed.) *Reduction, Modality and Time* (Cambridge: Cambridge University Press), 143-188.

[116] McGinn, C. (2001) *Logical Properties*. New York: Oxford University Press.

[117] Mena, R. (MS) 'Indeterminate Objects'

[118] McGinn, C. (2000) *Logical Properties*. New York: Oxford University Press.

[119] Nelson, J. (1970) *Noûs* 4 (3): 241-260.

[120] Nelson, M. (2007) 'Ways an Actualist might be' *Philosophical Studies* 133: 455-471.

[121] Noonan, H. (1993) 'Constitution is Identity' *Mind* 102 (405): 133-146.

[122] Noonan, H. (1997) 'Relative Identity' in B. Hale and C. Wright (eds.) *A Companion to the Philosophy of Language* (Oxford: Blackwell), 634-652.

[123] Oderberg, D. (1993) *The Metaphysics of Identity over Time*. New York: St. Martin's Press.

[124] Oliver, M. (1996) 'The Metaphysics of Properties', *Mind* 105 (417): 1-80.

[125] Over, D. E. (1989) 'Vague Objects and Identity' *Analysis* 49 (3): 97-99

[126] Parsons, T. and Woodruff, P. (1995) 'Worldly Indeterminacy of Identity' *Proceedings of the Aristetelian Society* 95: 171-191

[127] Parsons, T. (2000) *Indeterminate Identity*. New York: Oxford Universiy Press.

[128] Pelczar, M. and Rainsbury, J. (1998) 'The Indexical Character of Names' *Synthese* 114: 293-317.

[129] Perry, J. (1968) 'Can the self divide?' *Journal of Philosophy* 69 (16): 463-488.

[130] Perry, J. (1970) 'The Same F' *Philosophical Review* 79 (2): 181-200.

[131] Perry, J. (1978) 'Relative Identity and Relative Number' *Canadian Journal of Philosophy* 7 (1): 1-14.

[132] Perry, J. (2000) *The Problem of the Essential Indexical and Other Essays*. 2nd. ed. Chicago: Chicago University Press.

[133] Plantinga, A. (1974) *The Nature of Necessity*. Oxford: Clarendon Press.

[134] Putnam, H. (2005) *Ethics without Ontology*. New York: Oxford University Pres..

[135] Quine, W. V. O. (1943) 'Notes on Existence and Necesity' *Journal of Philosophy* 40 (5): 113-127.

[136] Quine, W. V. O. (1947) 'On the Problem of interpreting Modal Logic' *Journal of Symbolic Logic*

[137] Quine, W. V. O. (1948) 'On what there is' *Review of Metaphysics* 2 (5): 21-38.

[138] Quine, W. V. O. (1951) 'Two Dogmas of Empricism' *Philosophical Review* 60 (1): 20-43.

[139] Quine, W. V. O. (1953) 'Three grades of modal involvement' in *Proceedings of the XIth International Congress of Philosophy* 14, Brusels. Amsterdam: North Holland Publishing Co. Reprinted in Quine (1976) *The Ways of Paradox and other essays* (Cambridge, MA: Harvard University Press),

[140] Quine, W. V. O. (1960) *Word and Object*. Cambridge, MA: MIT Press.

[141] Quine, W. V. O. (1963) *From a Logical Point of View*. New York: Harper and Row.

[142] Quine, W. V. O (1969) *Ontological Relativity and other Essays*. New York City: Columbia University Press.

[143] Quine, W. V. O (1981) *Theories and Things*. Cambridge, MA: Harvard University Press.

[144] Quine, W. V. O. (1986) *Philosophy of Logic*. 2^{nd} ed. Cambridge, MA: Harvard University Press.

[145] Robinson, R. (1950) *Definition*. Oxford: Clarendon Press.

[146] Rosenberg, J. (1993) 'Comments on Peter Van Inwagen's *Mateial Beings*' *Philosophy and Phenomenological Research* 53 (3): 701-708

[147] Russell, B. (1910) 'Knowledge by Acquaintance and Knowledge by Description' *Proceedings of the Aristotelian Scoiety* 11: 108-128.

[148] Russell, B. (1923) 'Vagueness' *Australasian Journal of Philosophy and Psychology*, 1: 8492.

[149] Sainsbury, M. (1991) 'Is there Higher-Order Vagueness' *Philosophical Quarterly* 41 (163): 167-182.

[150] Salmon, N. (1981) *Reference and Essence* (1st ed.) Princeton: Princeton University Press.

[151] Salmon, N. (2005a) *Metaphysics, Mathematics and Meaning. Philosophical Papers. Vol. I* New York: Oxford University Press

[152] Savellos, E. (1990) 'On defining identity' *Notre Dame Journal of Formal Logic* 31 (3): 476-484.

[153] Sidelle, A. (1989) *Necesity, Essence and Individuation. A defense of Conventionalism*. Ithaca: Cornell University Press.

[154] Sidelle, A. (2002) 'On the Metaphysical Contingency of the Laws of Nature' in T. Szabó Gendler and J. Hawthorne (eds.) *Conceivability and Possibility* (New York: Oxford University Press), 309-336.

[155] Sider, T. (1999) 'Global Supervenience an Identity across Times and Worlds' *Philosophy and Phenomenological Research* 59 (4): 913-937

[156] Sider, T. (2001) *Four-Dimensionalism*. New York: Oxford University Press.

[157] Sider, T. (ms) 'Beyond the Humphrey Objection'

[158] Smart, J. J. C. (1959) 'Sensations and brain processes' *Philosophical Review* 68 (2): 141-156.

[159] Soames, S. (2002) *Beyond Rigidity*. New York: Oxford University Press.

[160] Soames, Scott (2006) 'The philosophical significance of the Kripkean necessary a posteriori' *Philosophical Issues* 16: 288-309.

[161] Stalnaker (1978) 'Assertion' in P. Cole (ed.) *Syntax and Semantics* 9 (Academic Press: New York), 315-332.

[162] Stalnaker, R. (1986) 'Counterparsts and Identity' *Midwest Studies in Philosophy* 9: 121-140

[163] Stalnaker, R. (1997) 'Reference and Necessity' in B. Hale and C. Wright (eds.) *A Companion to the Philosophy of Language* (Oxford: Blackwell), 534-554.

[164] Stanley, J. (1997) 'Names and rigid desiganation' in B. Hale and C. Wright (eds.) *A Companion to the Philosophy of Language* (Oxford: Blackwell), 555-585.

[165] Strawson, P. (1959) *Individuals*. London: Methuen.

[166] Thomason, R. (1982) 'Identity and Vagueness' *Philosophical Studies* 42 (3): 329-332.

[167] Tye, M. (1990) 'Vague objects' *Mind* 99 (536): 535-557.

[168] Van Inwagen (1990) *Material Beings*. Ithaca: Cornell University Press.

[169] Van Inwagen (2001) *Ontology, Identity and Modality*. New York: Cambridge University Press.

[170] Varzi, A. (2001) 'Vagueness in Geography' *Philosophy and Geography* 4 (1): 49-65.

[171] Wallace, J. (1965) 'Sortal predicates and quantification' *Journal of Philosophy* 62: 8-13.

[172] Wells, R. (1957) 'Frege's ontology' *The Review of Metaphysics* 4 (4): 537-573.

[173] Wiggins, D. (1967) *Identity and Spatio-Temporal Continuity*. Oxford: Basil Blackwell.

[174] Wiggins, D. (1997) 'Sortal concepts: a reply to Xu' *Mind and Language* 12 (3-4)): 413-421.

[175] Wiggins, D. (2001) *Sameness and Substance Renewed*. New York: Cambridge University Press.

[176] Williams, J. R. G. (2008) 'Multiple Actualities and Ontically Vague Identity' *Philosophical Quarterly* 58 (230): 134-154.

[177] Williamson, T. (1986) 'Criteria of Identity and the Axiom of Choice' *Journal of Philosophy* 83: 380-394.

[178] Williamson, T. (2002) 'Vagueness, Identity and Leibniz's Law' in A. Bottani *et al.* (eds.) *Individuals, Essence and Identity* (Dordrecht: Kluwer), 273-303.

[179] Wittgenstein, L. (1922) *Tractatus Logico-Philosophicus*. (trans. C. K. Ogden). London: Routledge and Kegan Paul.

[180] Wreen, M. (1998) 'Proper names and the necessity of identity statements' *Synthese* 114: 319-335.

[181] Wright, C. (1992) 'Is Higher-Order Vagueness Coherent' *Analysis* 52 (3): 129-139.

[182] Yablo, S. (1987) 'Identity, Essence and Indiscernibility' *Journal of Philosophy* 84 (6): 293-314.

[183] Yablo, S. (1998) 'Does Ontology rest on a Mistake?' *Proceedings of the Aristotelian Society. Supplementary Volumes* 72: 229-271.

[184] Xu, F. (1997) 'From Lot's wife to a pillar of salt: evidence that *physical object* is a sortal concept' *Mind and Language* 12 (3-4): 365-392.

[185] Zimmerman, D. (1998) "Criteria of Identity and the 'Identity Mystics"' *Erkenntnis* 48: 281-301: